DOCUMENT RESUME

ED 064 958 24 EM 010 208

TITLE	EIN Software Catalog; Entries, Volume II.
INSTITUTION	Interuniversity Communications Council (EDUCOM), Boston, Mass.
SPONS AGENCY	National Science Foundation, Washington, D.C.; Office of Education (DHEW), Washington, D.C.
BUREAU NO	BR-8-0697/Attachment No. 1
PUB DATE	[72]
GRANT	OEG-0-8-080697(007)
NOTE	755p.; See also EM 010 206, EM 010 207, EM 010 209, EM 010 210, EM 010 211, EM 010 212
EDRS PRICE	MF-$0.65 HC-$26.32
DESCRIPTORS	Behavioral Sciences; Biological Sciences; *Catalogs; Computer Oriented Programs; *Computer Programs; Earth Science; Educational Administration; Engineering; *Information Networks; Information Science; *Interinstitutional Cooperation; Library Science; Mathematics; Operations Research; Physical Sciences; *Program Descriptions; Statistics; Technology; Tests
IDENTIFIERS	*Educational Information Network; EIN

ABSTRACT

The EIN (Educational Information Network) is a non-profit operation which coordinates the sharing of educational computing resources. It is administered by EDUCOM and funded jointly by the U. S. Office of Education and the National Science Foundation. EIN maintains a group of contact personnel at member institutions to serve as a liaison between the institution and EIN. Through these persons items of software are offered for distribution. EIN also publishes a catalog of the software which is available through the network. The four-volume catalog contains an alphabetical listing of the participating EIN members that are represented in the catalog by program descriptions; descriptions of each computer facility listed and its general pricing algorithm; abstracts of available programs, subdivided into 13 areas of application; three indexes--by EIN number, by descriptive title, and by keyword; and complete descriptions of programs, including user instructions, samples of input and output, and cost estimates. Volume one contains a description of the EIN, facilities descriptions, abstracts, and indexes. This volume is one of three which contain the complete program descriptions. (JY)

ED 064958

EIN Software Catalog; Entries, Volume II.

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

EM 010 208

BR 8-0697
PA 56

EDUCOM EDUCATIONAL INFORMATION NETWORK

000 0076

DESCRIPTIVE TITLE	Algorithms for Analysis of Variance and Covariance of Incomplete Block and Lattice Designs
CALLING NAME	GAVIAL
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	K.E. Merritt Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

GAVIAL calculates analysis of variance and covariance for a wide variety of statistical problems occurring in the class of incomplete block designs. These include the classes of lattice design.

An algorithm, first proposed by Shah[1], is utilized to solve indirectly the set of reduced normal equations of treatment effects denoted by $A\hat{\tau} = Q$, where the matrix A is $(v \times v)$ of rank $v-1$, and $v \leq 200$ is the number of treatments in the design. Special properties of partially balanced incomplete block designs are exploited to obtain least squares estimates of treatment effects by solving a smaller set of equations denoted by $DZ = L$, where D is $(m \times m)$, and m is the number of distinct associate classes in the design. This method generally results in greater accuracy of estimation and increases the capability for solving larger problems.

GAVIAL utilizes a scan subroutine that simplifies, and for the most part minimizes, the effort required to specify problems to be analyzed. No particular statistical knowledge is required

continued

to use GAVIAL. Generally, a user only needs to know how many treatments or entries, plots, blocks, etc. are present in his experimental design.

REFERENCES

1. Shah, B.V., "A Generalization of Partially Balanced Incomplete Block Designs," Annals of Math. Stat., *30*, pp. 1041–1050, (1959).

USER INSTRUCTIONS

Detailed instructions are available from the GAVIAL Manual[1]. A few comments are included here to indicate the scope of the program.

GAVIAL may logically be divided into seven major components;

MAIN, LSCAN, INPUT, BIB, PBIB, LATTIC, and ORTHOG.

The subprogram MAIN initializes machine dependent parameters used in LSCAN and also acts as a supervisor for the other six major subroutines. LSCAN reads the problem specification supplied by the user, consisting of a set of keyword control cards which contain a value for each problem parameter.

INPUT is called next to read a data format card and the data supplied by the user. A check is made to determine if there is a common block size for all partially balanced incomplete block designs which do not exceed program limits specified for replicates, blocks and plots. If too many plots or too few plots were found in one or several blocks, an error message will be printed stating in which blocks a deficiency was found. When an error is detected, the program rejects the current problem and proceeds to the next problem. If no deficiency is found among the plots, the complete experimental design, as specified by subscripts associated with the data, will be printed. If transformations of data are requested, control is passed to subroutine TRANS which performs the transformation and then specifies that control be transferred to MAIN.

At this point control is passed to one of four algorithms; BIB, PBIB, LATTIC, or ORTHOG.

The BIB algorithm calculates analysis of variance for the class of balanced incomplete block designs and for the class of incomplete latin squares that are balanced incomplete block with respect to rows (blocks).

The PBIB algorithm produces an analysis for the class of balanced incomplete block and partially balanced incomplete block designs and for the class of incomplete latin square designs that are balanced with respect to columns and partially balanced with respect to rows (blocks).

The LATTIC algorithm produces an analysis for a special class of rectangular lattice design that does not satisfy the above

continued

definition of a partially balanced incomplete block design. This algorithm also calculates analysis for the class of triple square lattice designs that does satisfy the above definition.

The ORTHOG algorithm was designed primarily to calculate analysis for the class of orthogonal latin square designs having up to fourteen othogonal classifications. An observation may be classified may be classified by row, column, treatment, Latin letter, Greek letter, numeral, etc.

The following program parameter limits for specifying a design for analysis by any of the algorithms that use the particular parameters, are currently in effect.

Maximum number of treatments	200
Maximum number of blocks/design	100
Maximum number of replicates/design	100
Maximum number of associate classes/design	14
Maximum number of variates and covariates/design	9
Maximum number of data points/design	600
Maximum number of data points across designs	120,000
Maximum number of columns	200
Maximum number of classifications	14
Maximum number of levels per classification	200

REFERENCES

1. Merritt, K.E., *Stat. Num. Anal. and Data Process. Series, No. 11: GAVIAL Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

SAMPLE INPUT

```
TEST CASE FOR TRIANGULAR PBIB DESIGN 1965 AMS P1816 .
                                   ALGORITHM =2
                                   REPLICATES =  4
                                   BLOCKS      = 10
                                   TREATMENTS = 10
                                   PLOTS       = 4
                                   VARIATES    =  1
                                   OUTPUT = 5
                                   ENDS



            REPLICATE POSITION =    0
            BLOCK POSITION =        1
            COLUMN POSITION =       0
            TREAT POSITION =        2




            DATA

 1   3                  0.2310E 01
 1   5                  0.2810E 01
 1   6                  0.1650E 01
 1   9                  0.2580E 01
 2   2                  0.2510E 01
 2   7                  0.1410E 01
 2   8                  0.1900E 01
 2   9                  0.3060E 01
 3   1                  0.2890E 01
 3   3                  0.2290E 01
 3   8                  0.1950E 01
 3  10                  0.2040E 01
 4   4                  0.2540E 01
 4   6                  0.2090E 01
 4   7                  0.2360E 01
 4  10                  0.2030E 01
 5   1                  0.2280E 01
 5   4                  0.2810E 01
 5   9                  0.2200E 01
 5  10                  0.2070E 01
 6   2                  0.1770E 01
 6   3                  0.2490E 01
 6   4                  0.2310E 01
 6   7                  0.3020E 01
 7   1                  0.2720E 01
 7   2                  0.2290E 01
 7   6                  0.1570E 01
 7   8                  0.2600E 01
 8   4                  0.2810E 01
 8   5                  0.2990E 01
 8   6                  0.2280E 01
 8   8                  0.2440E 01
 9   1                  0.2540E 01
 9   3                  0.2440E 01
 9   5                  0.2230E 01
 9   7                  0.2120E 01
10   2                  0.1540E 01
10   5                  0.2870E 01
10   9                  0.2770E 01
10  10                  0.2090E 01
```

SAMPLE OUTPUT

```
                    EXPERIMENTAL PLAN

 BLOCK

     1        3    5    6    9
     2        2    7    8    9
     3        1    3    8   10
     4        4    6    7   10
     5        1    4    9   10
     6        2    3    4    7
     7        1    2    6    8
     8        4    5    6    8
     9        1    3    5    7
    10        2    5    9   10


                SPECIFICATION CF P.B.I.B. DESIGN

           ASSOCIATE CLASSES FOR TREATMENT   1

          CLASS      LAMDA  NO./CLASS         ASSOCIATE CLASSES

              1         12          1          1
              2          2          3          10  8  3
              3          1          6          9  7  6  5  4  2
           ASSOCIATE CLASSES FOR TREATMENT  10

          CLASS      LAMDA  NO./CLASS         ASSOCIATE CLASSES

              1         12          1          10
              2          2          3          9  4  1
              3          1          6          8  7  6  5  3  2

                P(I,J) ARRAY FOR TREAT   1   AND TREAT  10

     1        0   1   0
     2        1   0   2
     3        0   2   4
           ASSOCIATE CLASSES FOR TREATMENT   8

          CLASS      LAMDA  NO./CLASS         ASSOCIATE CLASSES

              1         12          1          8
              2          2          3          6  2  1
              3          1          6          10  9  7  5  4  3

                P(I,J) ARRAY FOR TREAT   1   AND TREAT   8

     1        0   1   0
     2        1   0   2
     3        0   2   4
```

continued

```
      ASSOCIATE CLASSES FOR TREATMENT    3

    CLASS      LAMDA  NO./CLASS           ASSOCIATE CLASSES

         1         12          1           3
         2          2          3           7  5  1
         3          1          6          10  9  8  6  4  2

             P(I,J) ARRAY FOR TREAT   1   AND TREAT   3

 1         0  1  0
 2         1  0  2
 3         0  2  4
      ASSOCIATE CLASSES FOR TREATMENT    9

    CLASS      LAMDA  NO./CLASS           ASSOCIATE CLASSES
         1         12          1           9
         2          2          3          10  5  2
         3          1          6           8  7  6  4  3  1

             P(I,J) ARRAY FOR TREAT   1   AND TREAT   9

 1         0  0  1
 2         0  1  2
 3         1  2  3
      ASSOCIATE CLASSES FOR TREATMENT    7

    CLASS      LAMDA  NO./CLASS           ASSOCIATE CLASSES

         1         12          1           7
         2          2          3           4  3  2
         3          1          6          10  9  8  6  5  1

             P(I,J) ARRAY FOR TREAT   1   AND TREAT   7

 1         0  0  1
 2         0  1  2
 3         1  2  3
      ASSOCIATE CLASSES FOR TREATMENT    6

    CLASS      LAMDA  NO./CLASS           ASSOCIATE CLASSES

         1         12          1           6
         2          2          3           8  5  4
         3          1          6          10  9  7  3  2  1

             P(I,J) ARRAY FOR TREAT   1   AND TREAT   6

 1         0  0  1
 2         0  1  2
 3         1  2  3
      ASSOCIATE CLASSES FOR TREATMENT    5

    CLASS      LAMDA  NO./CLASS           ASSOCIATE CLASSES

         1         12          1           5
         2          2          3           9  6  3
         3          1          6          10  8  7  4  2  1

             P(I,J) ARRAY FOR TREAT   1   AND TREAT   5

 1         0  0  1
 2         0  1  2
 3         1  2  3
```

continued

```
   ASSOCIATE CLASSES FOR TREATMENT   4

  CLASS     LAMDA  NO./CLASS         ASSOCIATE CLASSES

       1        12          1         4
       2         2          3        10  7  6
       3         1          6         9  8  5  3  2  1

         P(I,J) ARRAY FOR TREAT   1   AND TREAT   4

1          0  0  1
2          0  1  2
3          1  2  3

   ASSOCIATE CLASSES FOR TREATMENT   2

  CLASS     LAMDA  NO./CLASS         ASSOCIATE CLASSES

       1        12          1         2
       2         2          3         9  8  7
       3         1          6        10  6  5  4  3  1

         P(I,J) ARRAY FOR TREAT   1   AND TREAT   2

1          0  0  1
2          0  1  2
3          1  2  3


                                P.B.I.B. TOTALS AND MEANS FOR VARIATE  1

TREAT NO.        TREAT TOTAL    BLK TOT/TREAT          Q            QBAR    TREAT(INTRA-BLK)   TREAT(COMBINED)  MEAN (ADJ)

     1               10.4300          37.0400          1.1700          6.5933          0.3339            0.2199       2.5616
     2                8.1100          36.9200         -1.1200         -9.6494         -0.3362           -0.2937       2.0461
     3                9.5300          37.4400          0.1700          1.0771          0.1098            0.0076       2.3493
     4               10.4700          38.4900          0.8475          9.4210          0.1926            0.3302       2.6720
     5               10.9000          38.4700          1.2825         12.3565          0.3909            0.3821       2.7239
     6                7.5900          38.0700         -1.9275        -11.2151         -0.5645           -0.3756       1.9661
     7                8.9100          36.8200         -0.2950         -4.2985         -0.0941           -0.1329       2.2089
     8                8.8900          37.7500         -0.5475         -2.7984         -0.2121           -0.0609       2.2809
     9               10.6100          36.8600          1.3950          7.5168          0.4159            0.2425       2.5823
    10                8.2300          36.8200         -0.9750         -8.9965         -0.2275           -0.3156       2.0261

TOTALS               93.6698         374.6792          0.0000          0.0029         -0.0000           -0.0002       2.3417




                  BLOCK TUTALS

                      9.3500           8.8800          9.1700          9.0200          9.3600            9.5900       9.1800
                     10.5200           9.3300          9.2700
```

continued

```
                         2.8.1.8 ANALYSIS OF VARIANCE __1

   SOURCE               D.F.          SUM OF SQS.                      M.S.
BLOCKS IGNOR TREATS      9                0.456299                  0.050700
TREATS ELIM BLOCKS       9                3.482961                  0.386996
ERROR                   21                3.039622                  0.144744
TOTAL                   39                6.978882

* BLOCKS ELIM TREATS     9                0.754552                  0.083839
* TREATS IGNOR BLOCKS    9                3.184708                  0.353856

**TREAT(COMBINED)        9               21.745621                  2.416180
```

COST ESTIMATE

For the job listed on the Sample Input, the chargeable computer time at the current rate for the Iowa State University ($375./hr.) was $5.00 (ISU minimum); postage and handling included.

Charge to user = computer time + postage and handling + network overhead
= $5.00 + network overhead

CONTENTS—GAVIAL

pages

1– 2 Identification & Abstract
3– 4 User Instructions
5– 9 I/O
11 Cost—Contents

DESCRIPTIVE TITLE	An Algorithm for the Optimization of a Quadratic Form Subject to Linear Restraints
CALLING NAME	ZORILLA
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	D.J. Soults J.J. Zrubeck V.A. Sposito Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

ZORILLA will solve quadratic programming problems on the IBM 360 system. The program is composed of a number of subprograms; each is called by a procedure control card. The sequence of control cards defines the solution procedure.

A manual[1] is available to inform the user about the correct formulation for optimizing a quadratic form subject to linear restrictions, and to provide the user with a detailed explanation of how to use the program (order of data deck, key punching format, control cards, etc.).

The program can be used to minimize numerical problems. Keyword commands such as SCAN or MODEL serve as an aid in finding invalid input data or an incorrectly specified model. The program can scale poorly defined problems upon the use of the SCALE agendum card.

continued

ZORILLA will find the values for $x_1, x_2, \ldots, x_n$ which optimize

$$f(x) = p'x - \tfrac{1}{2}\, x'Cx$$

subject to the conditions that

$$Ax \leq b$$
$$x \geq 0$$

where p, C, A contain specified constants. The function f(x) will be referred to as the function which we are optimizing. C must be an n X n symmetric matrix. [Any quadratic form x'Bx may be expressed by a symmetric matrix $x'(\frac{B+B'}{2})x = \frac{1}{2}\, x'Cx$ where $C = (B+B')$.] It is also necessary for C to be a positive (or negative) semi-definite matrix when the objective function f(x) is required to be concave (or convex) in the maximization (or minimization) case.

The program requires a simplex tableau input in the form

$$\begin{array}{c@{\;}c} & \begin{array}{cc} n & 1 \end{array} \\ \begin{array}{c} n \\ m \end{array} & \begin{bmatrix} -C & -p \\ A & b \end{bmatrix} \end{array}$$

n = number of variables
m = number of restrictions.

The identity matrix need not be entered explicitly into the matrix. It is generated by the program.

The simplex method is utilized for transformation. The procedure for choosing the incoming and outgoing vectors is the procedure developed by Van de Panne and Whinston[2]. The method of transforming and updating vectors is the product form of the inverse based on the revised simplex method.

REFERENCES

1. Soults, D.J., Zrubeck, J.J., and Sposito, V.A., *Stat. Num. Anal. and Data Process. Series, No. 9: ZORILLA Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

2. Van de Panne, C., and Whinston, A., "Simplicial Methods for Quadratic Programming," Naval Research Logistics Quarterly, *11*, pp. 273–302, (1964).

USER INSTRUCTIONS

Detailed instructions are available from the ZORILLA Manual.[1]

REFERENCE

1. Soults, D.J., Zrubeck, J.J., and Sposito, V.A., *Stat. Num. Anal. and Data Process. Series, No. 9: ZORILLA Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

SAMPLE INPUT

```
INPUT
SCAN
ROW.ID
              ROW1
              ROW2
              ROW3
             +ROW4
             -ROW5
MATRIX
        COL1  ROW1      -4.0
        COL1  ROW2       2.0
        COL1  ROW4       1.0
        COL2  ROW1       2.0
        COL2  ROW2      -4.0
        COL2  ROW4       1.0
        COL2  ROW5       1.0
        COL3  ROW5      -1.0
FIRSTB
        ZZ    ROW1      -6.0
        ZZ    ROW4       2.0
        ZZ    ROW5       1.0
ENDATA
MAX..
OUTPUT
ENDJOB
```

SAMPLE OUTPUT

```
                                      ISU 360 Q.P.-DISK VERSION

 INVERT FREQ. PIVOT TOL.  PRICE TOL.  FEASB.TOL.  ARTFL.TOL.  ELEMT.TOL.,
     9999     .000100000 .000100000 .001000000 .000100000 .000001000




 NUMBER ROWS IN A     VARIABLES    RHS    DENSITY
                2             3      1         12

 AGENDUM TIME WAS   0.062 MINUTES.




           MAX..              PRICING FRACTION  1.000
                ITERATION ACCOUNTING
   ITR     ENTER  BASIS   LEAVE      ZJ-CJ S     B    NB    INFEAS

    1     10002     5    20005          -1     0    0       0.0         9  101        -1.00000
    2     30005     2    40002          -1     2    5       0.0        12  101        -1.00000
 SOLUTION FEASIBLE.


 ITERATION TIME WAS   0.045 MINUTES
```

continued

```
                                                          BASIS INVERSION.                  BASIS SAVED.
                 ITERATION ACCOUNTING
 ITR    ENTER  BASIS   LEAVE     ZJ-CJ S     B    NB     INFEAS
                                                                    INVERT TIME WAS    0.028 MINUTES.
SOLUTION FEASIBLE.

   3    10001    4     20004           1     0     0        0.0        14   101         0.0
   4    30004    1     40001           1     1     4        0.0        18   101         0.0
SOLUTION OPTIMAL.

AGENDUM TIME WAS    0.243 MINUTES




            OUTPUT


 OPTIMAL SOLUTION VALUES

 STRUCTURAL VARB.               OPT. LEVEL

  COL1    10001              0.100000000D 01
  COL2    10002              0.100000000D 01


  PRIMAL SLACKS                 OPT. LEVEL



  LAGRANGE MULT.                OPT. LEVEL

  ROW4    30004   1          0.400000000D 01
  ROW5    30005  -1         -0.600000000D 01


  LAGRANGE SLACK                OPT. LEVEL

  ROW3    40003              0.600000000D 01

       FUNCTIONAL VALUE IS   0.400000000D 01


  AGENDUM TIME WAS    0.043 MINUTES
```

COST ESTIMATE

For the job listed on the Sample Input, the chargeable computer time at the current rate for the Iowa State University ($375./hr.) was $5.00 (ISU minimum); postage and handling included.

Charge to user = computer time + postage and handling + network overhead
= $5.00 + network overhead

CONTENTS—ZORILLA

pages
1– 2 Identification & Abstract
3 User Instructions
5– 7 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	A Regression Model Building System
CALLING NAME	MOUFLON
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Bonnie L. Hanson Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. William J. Kennedy, Head, Statistical Numerical Analysis and Data Processing Section, Statistical Laboratory, Iowa State University, Ames, Iowa 50010 Tel.: (515) 294-2260

FUNCTIONAL ABSTRACT

MOUFLON provides fast and economical use of computational methods for model building in multiple linear regression. The basic model used is

$$y_i = \beta_0 x_{i0} + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + e_i$$

$$i = 1,2, \dots ,n$$

where x_{ij}'s are fixed, y_i's are observed values, and e_i's are assumed to be independent $N(0,\sigma^2)$. The values $x_{i0} \equiv 1$, $i = 1, \dots ,n$, are usually used. If we substitute $\beta_0 = \bar{y} - \beta_1 \bar{x}_1 - \dots - \beta_k \bar{x}_k$ into the original model, the result is

$$y_i - \bar{y} = \beta_1 (x_{i1} - \bar{x}_1) + \dots + \beta_k (x_{ik} - \bar{x}_k) + e_i$$

or

$$y_i' = \beta_1 x_{i1}' + \dots + \beta_k x_{ik}' + e_i ,$$

which is called the reduced model. The usual procedure solves the reduced normal equations associated with the above model and obtains the constant term by means of the relation

continued

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x}_1 - \ldots - \hat{\beta}_k\bar{x}_k \quad .$$

It may be helpful to note that the coefficient matrix corresponding to the reduced normal equations is in fact the matrix of corrected sums of squares and cross products. This coefficient matrix will be called X'X in the future. The following notation is used throughout.

k = total number of independent variables available for inclusion in the regression,

p = number of variables currently included in the regression,

RSS_p = residual sum of squares associated with a p–variate regression

RMS = residual mean sqaure,

REGSS = sum of squares due to regression,

REGMS = regression mean square,

$\hat{\beta}_i^2/c_{ii}$ = reduction in REGSS when ith variable is removed from the regression ($\hat{\beta}_i$ is the estimator of β_i and c_{ii} is the ith diagonal element of $(X'X)^{-1}$

Methods

Four methods are available for selecting the optimal regression.

Hocking and Leslie's Use of the C_p Statistic

Hocking and Leslie's procedure[1] is based on earlier work done by C.L. Mallows.[2,3] Mallows suggested that the selection of a 'good' subset of independent variables in a multiple linear regression be based on the standardized total squared error. He developed the C_p statistic as an estimate of this quantity.

$$C_p = \frac{RSS_p}{\hat{\sigma}^2} - (n - 2p)$$

where $\hat{\sigma}^2$ is the residual mean square obtained by fitting the full model and is used as an estimate of σ^2. Mallows showed that models with small bias tend to have C_p's almost equal to p, and he defined these to be 'good' models. Thus, a graph of C_p versus p will show which of the subsets of independent variates are 'good'. However, this method requires the computation of all regressions.

continued

Hocking and Leslie developed an efficient procedure to find the 'best' (in terms of minimum residual sum of squares) subset or subsets of independent variates. Their approach to the problem is concerned with which subset of r variates, $r = k - p$, should be removed from the regression. It is necessary to note that a minimum reduction in the regression sum of squares due to removing a set of r variables implies that the associated p–variate regression has minimum residual sum of squares. Hocking and Leslie state that an equivalent form of the C_p statistic is

$$C_p = \frac{\text{Red}_p}{\hat{\sigma}^2} + (2p - k)$$

where Red_p is the reduction in regression sum of squares due to removing a set of r variables, where $r = k - p$.

The above equations relate to the reduced model

$$y = \beta_1 + \beta_2 x_2 + \ldots + \beta_k x_k + e$$

where β_1 is the constant term. After adjustment for the notation and model used by the program, the C_p statistic is represented as

$$C_p = \frac{\text{Red}_p}{\hat{\sigma}^2} + (2p - k + 1).$$

The first step of Hocking and Leslie's procedure is to compute the k univariate reductions, θ_i (i.e. θ_i = reduction due to removing the ith variate) and to rank them so that $\theta_1 \leq \theta_2 \leq \cdots \leq \theta_k$. At the same time the variables are relabeled according to the order on the θ_i's. Thus, the removal of the first variable (according to the above labeling) leaves the subset of size $k - 1$ with minimum residual sum of squares among the set of possible subsets of size $k - 1$.

Hocking and Leslie's procedure is based on the following property.

> If the reduction in the regression sum of squares due to eliminating any set of variables for which j is the maximum subscript is not greater than θ_{j+1}, then no subset including any variables with subscripts greater than j can result in a smaller reduction.

The sequential procedure for a given $p = 1, \ldots, k - 2$ (which determine $r = 2, \ldots, k - 1$) is as follows.

continued

Step 1 Compute the reduction R due to removing variables 1,2,...,r (according to above labeling). If $R \le \theta_{r+1}$, procedure terminates and regression consisting of variables r + 1,...,k is taken to be the 'best' p–variate model. If $R > \theta_{r+1}$, go to Step 2.

Step 2 Include variable r + 1 in the set of variates which are candidates for removal. Then compute reductions for all subsets of size r which include variate r + 1. Find the smallest of all reductions calculated so far (the reduction found in Step 1 is included in the set). If the minimum reduction is not larger than θ_{r+2}, the procedure terminates and the p–variate set corresponding to the minimum reduction is taken to be the 'best' regression of size p. If the minimum reduction exceeds θ_{r+2}, go to Step 3.

Step 3 Add variable r + 2 to the set of variates which are candidates for removal. Compute the reductions for all subsets of size r which include variate r + 2. Find the smallest of all reductions calculated so far (the reductions found in Steps 1 and 2 are included in the set). If the minimum reduction is not larger than θ_{r+3}, the procedure terminates and the p–variate set corresponding to the minimum reduction is taken to be the 'best' regression of size p. If the minimum reduction exceeds θ_{r+3}, go to Step 4.

⋮

The procedure continues through as many steps as are necessary to find the subset with minimum residual sum of squares. After the 'best' subset of size p has been found the value of p is decreased by 1 (r increased by 1) and the procedure is executed once again to find the 'best' subset of size p - 1. Once the 'best' subset of size p' (specified by user) is found, the procedure terminates.

Forward Selection

This method[4] assumes the user has a predetermined order for his independent variate. The variates with highest degree of 'importance' for inclusion in the model must be placed first in the ordering and followed by those of lesser 'importance'.

Often, the user will have a basic set of r, $1 \le r \le k$, independent variates which are always needed to predict values of the dependent variable y, and a set of k - r variates which might be of value in the prediction equation. He thus wishes

continued

to find out which of the set of k - r variates should be used to predict y. The forward selection procedure begins with the basic set of r variates and proceeds to test each successive variate by means of the following sequential procedure.

Step 1 Test $H_1 : \beta_{r+1} = 0$
Acceptance implies procedure terminates and the basic set of r variates is used to predict y. Rejection implies procedure continues to Step 2 and the basic set of r variates plus variable r + 1 are included in the model.

Step 2 Test $H_2 : \beta_{r+2} = 0$
Acceptance implies procedure terminates and variates $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Rejection implies procedure continues to Step 3 and variates $x_1, x_2, \ldots, x_{r+1}, x_{r+2}$ are used to predict y.

Step 3 Test $H_3 : \beta_{r+3} = 0$

.
.
.

The basic test of $H_i : \beta_{r+i} = 0$ is to reject if $u > F_\alpha$ where

$$u = \frac{\hat{\beta}_i^2 / c_{ii}}{RMS}$$

F_α is specified by the user.

Sequential Deletion

This procedure assumes as in Forward Selection[4] that the independent variates' order begins with the 'most important' and ends with the 'least important'. Again, the user specifies a basic set r, $1 \leq r < k$, of independent variables needed to predict values of $\bar{y}$. The sequential deletion procedure begins with the full k-variate model and tests each preceding variate by means of the following sequential procedure.

Step 1 Test $H_0 : \beta_k = 0$
Rejection implies procedure terminates and all k independent variables are used to predict y. Acceptance implies procedure continues to Step 2 and the kth variate is removed from the model.

Step 2 Test $H_1 : \beta_{k-1} = 0$
Rejection implies procedure terminates and $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Acceptance implies procedure continues to Step 3 and the (k-1)st variate is removed from the model.

continued

Step 3 Test $H_2 : \beta_{k-2} = 0$

⋮

Step k - r Test $H_{k-r-1} : \beta_{r+1} = 0$
Rejection implies procedure terminates and $x_1, x_2, \ldots, x_{r+1}$ are used to predict y. Acceptance implies procedure terminates but x_{r+1} is removed from the model and only $x_1, x_2, \ldots, x_r$ are used to predict y. Again the test of $H_{k-i} : \beta_i = 0$ is to reject if $u > F_\alpha$ where

$$u = \frac{\hat{\beta}_i^2/c_{ii}}{\text{RMS}} .$$

F_α is specified by the user.

Stepwise Regression

Before we look at the stepwise procedure[5] let us consider a general case. First, let X_1 denote the set of variates $x_1, \ldots, x_p$ which are currently in the model and let X_2 denote the set of variates $x_{p+1}, \ldots, x_k$.

Next, define the sample partial correlation coefficient of the dependent variable y and one of the independent variates from the set X_2, (say x_{p+i}), to be the simple correlation coefficient of y^* and x^*_{p+i}. y^* is the set of residuals resulting from the regression of y on $x_1, x_2, \ldots, x_p$ and x^*_{p+i} is the set of residuals resulting from the regression of x_{p+i} on $x_1, x_2, \ldots, x_p$. Let the sample partial correlation coefficient of y and x_{p+i} be denoted by r_{p+i}.

Now consider the case of transferring one of the variates, x_{p+j}, from set X_2 to set X_1 , (i.e. including x_{p+j} in the regression equation). Let RSS_{p+j} denote the residual sum of squares for the regression of y on $x_1, x_2, \ldots, x_p, x_{p+j}$. Note that the value of j can be $1, 2, \ldots, k-p$.

Let x_{p+i} be 'the' variable transferred to the set X_1 where i satisfies

$$RSS_{p+i} \leq RSS_{p+j} \qquad j = 1, 2, \ldots, k-p$$

This is equivalent to

$$r^2_{p+i} \geq r^2_{p+j} \qquad j = 1, 2, \ldots, k-p$$

since it can be shown that

$$RSS_{p+j} = RSS_p(1 - r^2_{p+j}) .$$

continued

Thus, the selection of the largest r^2_{p+j} for $j = 1,2,\ldots,k-p$, minimizes RSS_{p+j}.

Now consider the case of transferring one of the independent variables x_j from the set X_1 to the set X_2 (i.e. removing it from the regression). Let RSS_j denote the residual sum of squares for the regression of y on $x_1, x_2, \ldots, x_{j-1}, x_{j+1}, \ldots, x_p$. It can be shown that

$$RSS_j = RSS_p + \frac{\hat{\beta}^2_j}{c_{jj}} \ .$$

Let x_i be 'the' variable transferred to the set X_2 where i satisfies

$$\frac{\hat{\beta}^2_i}{c_{ii}} \leq \frac{\hat{\beta}^2_j}{c_{jj}} \qquad j = 1,2,\ldots,p$$

or

$$(t^2_c)_i \leq (t^2_c)_j \qquad j = 1,2,\ldots,p$$

where $(t^2_c)_i = \frac{\hat{\beta}^2_i / c_{ii}}{RMS}$. Thus, the selection of the smallest $(t^2{}_c)_j$ for $j = 1,2,\ldots,p$ minimizes RSS_j.

The above considerations give way to the stepwise regression procedure, which consists of two alternating steps and examination of termination criteria after each step. The procedure terminates when any one of the following criteria is encountered.

1. There is no variable to enter and no variable to remove.
2. The procedure dictates that the same variable be entered and removed successively. This can be corrected by changing the F levels if the user so wishes.
3. The total number of steps executed reaches the maximum number of steps specified by the user.

The procedure begins with Step 1 and no variables entered in the model.

continued

Step 1 Enter variable i into the regression if i satisfies

$$r^2_{p+i} \geq r^2_{p+j} \quad j = 1,2,\ldots,k\text{-}p \quad \text{and} \quad (t^2_c)_{p+i} = \frac{(n-p-2)r^2_{p+i}}{(1-r^2_{p+i})} > F_{in} ,$$

where F_{in} is the F level to enter a variable and is specified by the user.

The termination criteria are now checked. If any one of the three criteria is satisfied, the program stops computations. If none of the criteria are satisfied, the program continues to Step 2.

Step 2 Remove variable i from the regression if i satisfies

$$(t^2_c)_i \leq (t^2_c)_j \quad j = 1,2,\ldots,p \quad \text{and} \quad (t^2_c)_i < F_{out} ,$$

where F_{out} is the F level to remove a variable and is specified by the user.

The termination criteria are now checked. If any one of the three criteria are satisfied, the program stops. If none of the criteria are satisfied, the program returns to Step 1.

REFERENCES

1. Hocking, R.R., and Leslie, R.N., "Selection of the Best Subset in Regression Analysis," Technometrics, *9*, 1967, pp. 531–540.
2. Mallows, C.L., *Choosing Variables in a Linear Regression: a Graphical Aid.* Paper presented at the Cent. Reg. Meeting of the Inst. of Math. Stat., Manhattan, Kansas, 1964.
3. Mallows, C.L., *Choosing a Subset Regression.* Paper presented at the Joint Stat. Meeting, Los Angeles, Cal., 1966.
4. Larson, H.J., *Sequential Model Building for Prediction in Regression Analysis*, unpublished Ph.D. thesis, (Iowa State University Library, Ames, Iowa, 1960).
5. Hemmerle, W.J., *Statistical Computations on a Digital Computer*, (Blaisdell Publishing Company, a Division of Ginn and Company, Waltham, Mass., 1967).

continued

6. Garside, M.J., "The Best Subset in Multiple Regression Analysis," Appl. Stat. J. of the Roy. Stat. Soc., *14*, Series C, 1965.

7. Hanson, B.L., *Stat. Num. Anal. and Data Process. Series, No. 12: MOUFLON Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

USER INSTRUCTIONS

Detailed instructions are available from the MOUFLON Manual.[1]

REFERENCE

1. Hanson, B.L., *Stat. Num. Anal. and Data Process. Series, No. 12: MOUFLON Reference Manual*, (Statistical Laboratory, Iowa State University, Ames, Iowa, 1969). Available from the Statistical Laboratory at a cost of $0.75.

SAMPLE INPUT

```
(System Control Cards)
JOBS=     1
BEGIN EXECUTION    4  13    1    1    1
TRANS
(5F12.5)
     7.00000    26.00000     6.00000    60.00000    78.50000
     1.00000    29.00000    15.00000    52.00000    74.30000
    11.00000    56.00000     8.00000    20.00000   104.30000
    11.00000    31.00000     8.00000    47.00000    87.60000
     7.00000    52.00000     6.00000    33.00000    95.90000
    11.00000    55.00000     9.00000    22.00000   109.20000
     3.00000    71.00000    17.00000     6.00000   102.70000
     1.00000    31.00000    22.00000    44.00000    72.50000
     2.00000    54.00000    18.00000    22.00000    93.10000
    21.00000    47.00000     4.00000    26.00000   115.90000
     1.00000    40.00000    23.00000    34.00000    83.80000
    11.00000    66.00000     9.00000    12.00000   113.30000
    10.00000    68.00000     8.00000    12.00000   109.40000
HOCKING       2    1
FORWARD       2        13.000
BACKWARD      1        13.000
STEPWISE               43.0003.000
ENDJOB        1
/*
```

SAMPLE OUTPUT

```
LIST OF OBSERVATIONS INPUT
    1
      7.00000     26.00000      6.00000     60.00000     78.50000
    2
      1.00000     29.00000     15.00000     52.00000     74.30000
    3
     11.00000     56.00000      8.00000     20.00000    104.30000
    4
     11.00000     31.00000      8.00000     47.00000     87.60000
    5
      7.00000     52.00000      6.00000     33.00000     95.90000
    6
     11.00000     55.00000      9.00000     22.00000    109.20000
    7
      3.00000     71.00000     17.00000      6.00000    102.70000
    8
      1.00000     31.00000     22.00000     44.00000     72.50000
    9
      2.00000     54.00000     18.00000     22.00000     93.10000
   10
     21.00000     47.00000      4.00000     26.00000    115.90000
   11
      1.00000     40.00000     23.00000     34.00000     83.80000
   12
     11.00000     66.00000      9.00000     12.00000    113.30000
   13
     10.00000     68.00000      8.00000     12.00000    109.40000
```

```
**********************************************************************************************************************

                                        SELECTION OF SUBSETS BY USE OF C SUB P STATISTIC

                                        ALGORITHM BY HOCKING-LESLIE --TECHNO. NOV. 1967

**********************************************************************************************************************


                        PROGRAM SELECTS BEST (IN SENSE OF MINIMUM RESIDUAL SS) THREE R-VARIATE
                        SUBSETS (R SPECIFIED BY USER) TO REMOVE FROM REGRESSION.

                        PROGRAM CONTINUES BY SELECTING BEST (R+1)-VARIATE SUBSETS TO REMOVE
                        FROM REGRESSION


                        . . .


                        PROGRAM STOPS EITHER WHEN BEST THREE R"-VARIATE SUBSETS (R" SPECIFIED BY
                        USER) HAVE BEEN SELECTED OR WHEN THERE IS NO POINT IN LOOKING AT SMALLER
                        VALUES OF P (NO. OF VARIABLES CURRENTLY IN REGRESSION).
```

```
                        * * * * * *

                        FULL MODEL

                        * * * * * *

                        INDEPENDENT VARIABLES  *  1  *  2  *  3  *  4  *
```

continued

```
                                        * * * * * * * * * * *.

                                        UNIVARIATE REDUCTIONS

                                        * * * * * * * * * * *


                               VAR NO              REDUCTION                     C SUB P

                                 3                  0.10909                      3.0182
                                 4                  0.24697                      3.0413
                                 2                  2.9725                       3.4968
                                 1                  25.951                       7.3375
```

```
                                        * * * * * * * * * * * * * * *

                                        3   BEST EQUATIONS FOLLOW

                                        SIZE OF SUBSET REMOVED =  2

                                        * * * * * * * * * * * * * * *



          SET REMOVED   *  3  *  4  *
              (CALL THIS SET B2)
              ((B1,B2) IS TOTAL SET)


                                        ANALYSIS OF VARIANCE

VARIATION DUE TO         DF          SUM OF SQUARES                    MEAN SQUARE                   F RATIO

 TOTAL                   12          0.271576E 04

 (B1,B2)                  4          0.266790E 04

 (B1)                     2          0.265786E 04

 (B2 ADJ FOR B1)          2          0.100408E 02                 0.502039E 01                  0.8391

 RESIDUAL                 8          0.478636E 02                 0.598295E 01


      MULTIPLE R SQUARE =0.978678


      C SUB P =     0.267823E 01
```

continued

```
          SET REMOVED  *  2  *  3  *
             (CALL THIS SET B2)
             ((B1,B2) IS TOTAL SET)


                                            ANALYSIS OF VARIANCE

VARIATION DUE TO          DF               SUM OF SQUARES                       MEAN SQUARE                F RATIO
 TOTAL                    12                0.271576E 04
 (B1,B2)                   4                0.266790E 04
 (B1)                      2                0.264100E 04
 (B2 ADJ FOR B1)           2                0.268984E 02                     0.134492E 02                 2.2479
 RESIDUAL                  8                0.478636E 02                     0.598295E 01


        MULTIPLE R SQUARE =0.972471


        C SUB P =     0.549584E 01
```

```
          SET REMOVED  *  2  *  4  *
             (CALL THIS SET B2)
             ((B1,B2) IS TOTAL SET)


                                            ANALYSIS OF VARIANCE

VARIATION DUE TO          DF               SUM OF SQUARES                       MEAN SQUARE                F RATIO
 TOTAL                    12                0.271576E 04
 (B1,B2)                   4                0.266790E 04
 (B1)                      2                0.148869E 04
 (B2 ADJ FOR B1)           2                0.117921E 04                     0.589604E 03                98.5473
 RESIDUAL                  8                0.478636E 02                     0.598295E 01


        MULTIPLE R SQUARE =0.548167


        C SUB P =     0.198095E 03



                          ALGORITHM TERMINATES NORMALLY
```

continued

```
****************************************************************************************************

                                       FORWARD SELECTION


                                       F TEST VALUE =    3.00000



****************************************************************************************************

                                       REGRESSION ANALYSIS

****************************************************************************************************

INDEPENDENT VARIABLES  *  1  *  2  *

                                       ANALYSIS OF VARIANCE

VARIATION DUE TO      DF            SUM OF SQUARES                    MEAN SQUARE                 F RATIO
      TOTAL           12             0.271576E 04
      REGRESSION       2             0.265786E 04              0.132893E 04             0.2295E 03
      RESIDUAL        10             0.579045E 02              0.579045E 01

      MULTIPLE R SQUARE =    0.97868


                        INFORMATION CONCERNING ESTIMATED COEFFICIENTS IN PREDICTION EQUATION

                        VAR NO.             COEFFICIENT              T VALJE            STANDARD ERROR
                             1               1.46831                 12.1047              0.121301
                             2              0.662250                 14.4424              0.458547E-01
                      INTERCEPT              52.5773                 22.9980              2.28617

                                       OBSERVED VS PREDICTED VALUES

                         OBSN NO.           OBSERVED                 PREDICTED            DIFFEREVCE
                              1              78.5000                  80.0740             -1.57400
                              2              74.3000                  73.2509              1.04908
                              3              104.300                  105.815             -1.51474
                              4              87.6000                  89.2585             -1.65848
                              5              95.9000                  97.2925             -1.39251
                              6              109.200                  105.152              4.04751
                              7              102.700                  104.002             -1.30205
                              8              72.5000                  74.5754             -2.07542
                              9              93.1000                  91.2755              1.82451
                             10              115.900                  114.538              1.36246
                             11              83.8000                  80.5357              3.26433
                             12              113.300                  112.437             0.862756
                             13              109.400                  112.293             -2.89344
                                                                                       -----------
                                                                SUM OF RESIDUALS     0.387246D-12



                        FORWARD SELECTION ALGORITHM TERMINATES NORMALLY
```

continued

```
****************************************************************************************************

                                    SEQUENTIAL DELETION

                                    F TEST VALUE =    3.00000



****************************************************************************************************

                                    REGRESSION ANALYSIS

****************************************************************************************************

INDEPENDENT VARIABLES  *  1  *  2  *  3  *  4  *

                                    ANALYSIS OF VARIANCE

VARIATION DUE TO      DF             SUM OF SQUARES                     MEAN SQUARE                  F RATIO
      TOTAL           12              0.271576E 04
      REGRESSION       4              0.266790E 04                     0.666975E 03                0.1115E 03
      RESIDUAL         8              0.478636E 02                     0.598295E 01

      MULTIPLE R SQUARE =    0.98238


                INFORMATION CONCERNING ESTIMATED COEFFICIENTS IN PREDICTION EQUATION

                VAR NO.              COEFFICIENT             T VALUE          STANDARD ERROR
                      1                1.55110               2.0827             0.744770
                      2                0.510168              0.7049             0.723788
                      3                0.101909              0.1350             0.754709
                      4               -0.144061             -0.2032             0.709052
              INTERCEPT               62.4054                0.8906            70.0709

****************************************************************************************************

                                    REGRESSION ANALYSIS

****************************************************************************************************

INDEPENDENT VARIABLES  *  1  *  2  *  3  *

                                    ANALYSIS OF VARIANCE

VARIATION DUE TO      DF             SUM OF SQUARES                     MEAN SQUARE                  F RATIO
      TOTAL           12              0.271576E 04
      REGRESSION       3              0.266765E 04                     0.889217E 03                0.1663E 03
      RESIDUAL         9              0.481106E 02                     0.534562E 01

      MULTIPLE R SQUARE =    0.98228




                INFORMATION CONCERNING ESTIMATED COEFFICIENTS IN PREDICTION EQUATION

                VAR NO.              COEFFICIENT             T VALUE          STANDARD ERROR
                      1                1.69589               8.2895             0.204582
                      2                0.656915             14.8508             0.442342E-01
                      3                0.250018              1.3536             0.184711
              INTERCEPT               48.1936               12.3153             3.91330
```

continued

```
                              OBSERVED VS PREDICTED VALUES

      OBSN NO.                OBSERVED               PREDICTED             DIFFERENCE
          1                   78.5000                78.6448              -0.144758
          2                   74.3000                72.6903                1.60968
          3                   104.300                105.636               -1.33580
          4                   87.6000                89.2129               -1.61293
          5                   95.9000                95.7245               0.175455
          6                   109.200                105.229                3.97110
          7                   102.700                104.173               -1.47256
          8                   72.5000                75.7543               -3.25427
          9                   93.1000                91.5591                1.54087
         10                   115.900                115.682               0.217602
         11                   83.8000                81.9165                1.88348
         12                   113.300                112.455               0.845033
         13                   109.400                111.823               -2.42289
                                                                       -----------
                                                SUM OF RESIDUALS     0.369482D-12




      SEQUENTIAL DELETION ALGORITHM TERMINATES NORMALLY
```

```
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *


                                     STEPWISE REGRESSION




                 F LEVEL TO ENTER =    3.00000

                 F LEVEL TO REMOVE =   3.00000




* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *


STEP NUMBER       1
VARIABLE ENTERED       4

MULTIPLE R SQUARE             0.6745
STD. ERROR OF EST.            8.9639

ANALYSIS OF VARIANCE

                             DF      SUM OF SQUARES     MEAN SQUARE       F RATIO
              REGRESSION      1          1831.8960      1831.8960         22.7985
              RESIDUAL       11           883.8669        80.3515
```

continued

```
VARIABLES IN EQUATION

          VARIABLE            COEFFICIENT             STD. ERROR             F TO REMOVE

             4               -0.73816                 0.15460                22.799

             (CONSTANT       117.56793)


VARIABLES NOT IN EQUATION

          VARIABLE            PARTIAL CORR.            F TO ENTER

             1                0.95677                 108.22
             2                0.13021                 0.17248
             3               -0.89508                 40.295



 * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *


STEP NUMBER     2
VARIABLE ENTERED      1

MULTIPLE R SQUARE            0.9725
STD. ERROR OF EST.           2.7343

ANALYSIS OF VARIANCE

                               DF     SUM OF SQUARES     MEAN SQUARE     F RATIO
          REGRESSION            2          2641.0007       1320.5002    176.6269
          RESIDUAL             10            74.7621          7.4762


VARIABLES IN EQUATION

          VARIABLE             COEFFICIENT             STD. ERROR             F TO REMOVE

             1                  1.4400                 0.13842                108.22
             4                 -0.61395                0.48645D-01            159.30

             (CONSTANT      103.09738)


VARIABLES NOT IN EQUATION

          VARIABLE             PARTIAL CORR.            F TO ENTER

             2                  0.59861                5.0259
             3                 -0.56571                4.2358


 * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
```

continued

```
STEP NUMBER    3
VARIABLE ENTERED     2

MULTIPLE R SQUARE           0.9823
STD. ERROR OF EST.          2.3087

ANALYSIS OF VARIANCE

                         DF     SUM OF SQUARES      MEAN SQUARE     F RATIO
          REGRESSION     3           2667.7900      889.2632       166.8316
          RESIDUAL       9             47.9727        5.3303


VARIABLES IN EQUATION

          VARIABLE             COEFFICIENT          STD. ERROR           F TO REMOVE

              1                  1.4519             0.11700              154.01
              2                 0.41611             0.18561              5.0259
              4                -0.23654             0.17329              1.8633

             (CONSTANT         71.64831)


VARIABLES NOT IN EQUATION

          VARIABLE             PARTIAL CORR.             F TO ENTER

              3                0.476860-01          0.182330-01



 * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *



STEP NUMBER    4
VARIABLE REMOVED     4

MULTIPLE R SQUARE           0.9787
STD. ERROR OF EST.          2.4063

ANALYSIS OF VARIANCE

                         DF     SUM OF SQUARES      MEAN SQUARE     F RATIO
          REGRESSION     2           2657.8584     1328.9292       229.5037
          RESIDUAL      10             57.9045        5.7904


VARIABLES IN EQUATION

          VARIABLE             COEFFICIENT          STD. ERROR           F TO REMOVE

              1                  1.4683             0.12130              146.52
              2                 0.66225             0.458550-01          208.58

             (CONSTANT         52.57735)


VARIABLES NOT IN EQUATION

          VARIABLE             PARTIAL CORR.             F TO ENTER

              3                  0.41126             1.8321
              4                 -0.41415             1.8633



                    STEPWISE ALGORITHM TERMINATED NORMALLY AFTER THIS STEP

                    THE REGRESSION MODEL IS AS GIVEN ABOVE
```

COST ESTIMATE

For the job listed on the Sample Input the total running time was 2.5 seconds. At the current rate for the Iowa State University ($375./hr.) the chargeable computer time was $5.00 (ISU minimum); postage and handling included.

Charge to user = computer time + postage and handling + network overhead
= $5.00 + network overhead

CONTENTS—MOUFLON

pages

1– 9 Identification & Abstract
11 User Instructions
13–21 I/O
23 Cost—Contents

DESCRIPTIVE TITLE	TEXT360: A System for Producing Manuals
CALLING NAME	TEXT360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	S.L. Reed, International Business Machines
LANGUAGE	PL/I and Assembly Language
COMPUTER	IBM/360 under OS
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	J. Philip Miller, Computing Facilities, Box 1152, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TEXT360 is a text-processing system with data-entry, data-updating, and page-formatting capabilities. The system, which runs under Operating System/360, consists of a main processor and several peripheral programs. Input to the system is free-form and is produced on the IBM 029 Card Punch. Output is camera-ready and is produced on the IBM 1403 Printer. The formatting capabilities of TEXT360 permit the insertion, deletion, and replacement of characters, words, lines, and groups of lines. In addition, blocks of text can be moved from one part of a document to another. One-column and two-column page format can be produced. Routine functions include hyphenation, line justification, column heading, and indentations. More complex functions include the generation of horizontal and vertical ruling for tables and figures. The program also allows the user to specify that related material (e.g., a table) is to be kept together, i.e., not split between columns or pages.

The four phases of the TEXT360 Formatting Processor, the TEXT360 Spelling Dictionary Update Program, the TEXT90-to-TEXT360 Master File Conversion Program and the Print/Punch Utility Program are

continued

written in the PL/1 Language, supplemented by four small assembler-language routines used for character-set mapping. The TEXT360 Prescan and Peripheral Print Programs are written in System/360 assembler language.

TEXT360 is essentially the program used by IBM for the production of the SRL manuals.

REFERENCES

Reed, S.L., *TEXT360*, Contributed Program Library Documentation, Program 360D-29.4.001 (White Plains, N. Y.: IBM, 1967).

USER INSTRUCTIONS

See Sample Input and reference listed below.

REFERENCES

Reed, S.L., *TEXT360*, Contributed Program Library Documentation, Program 360D-29.4.001 (White Plains, N. Y.: IBM, 1967).

SAMPLE INPUT AND OUTPUT

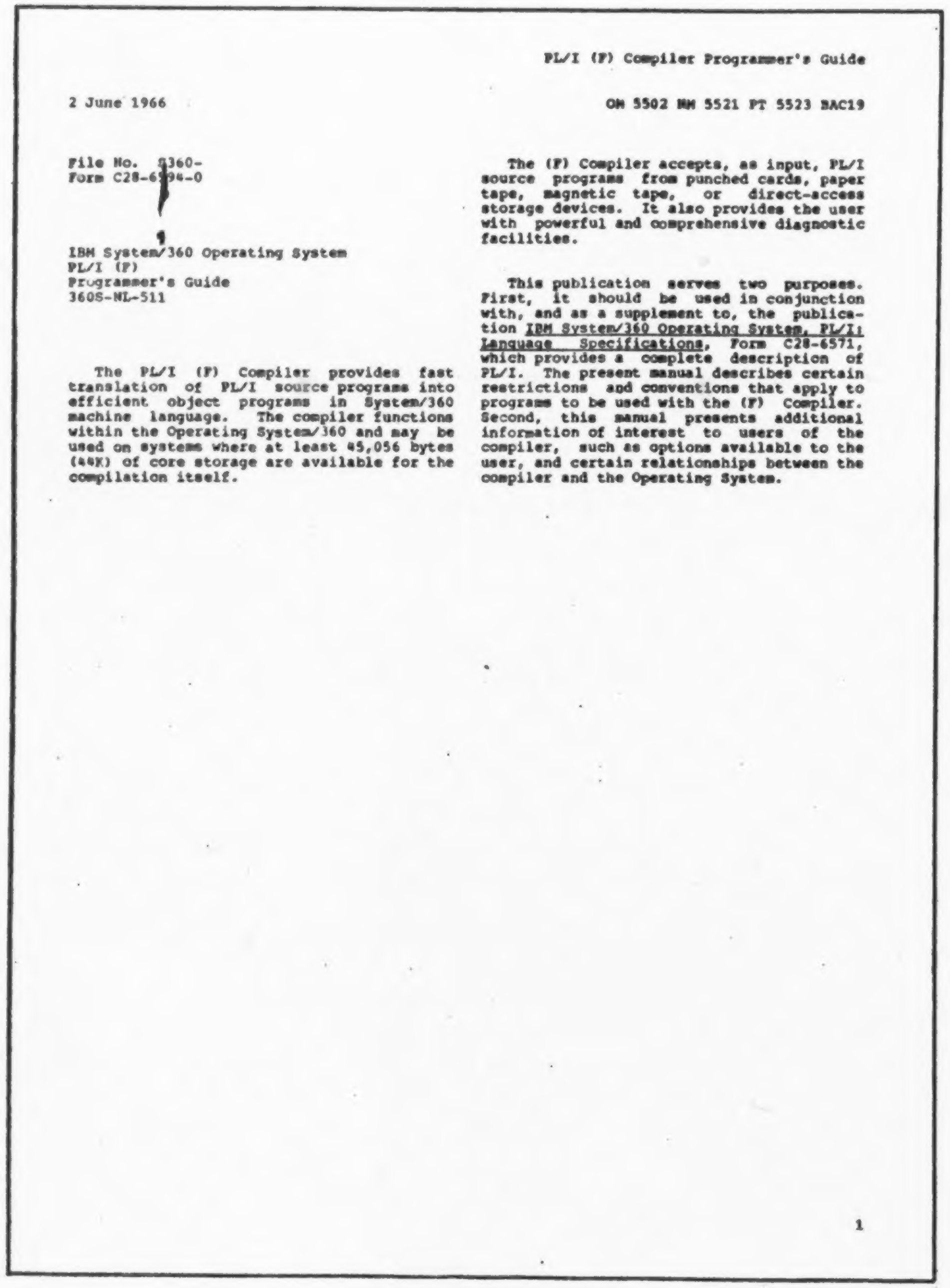
This sample contains three coding examples. The examples depict the coding required to initially format text, to update text, and to format a table. Each example consists of two parts: the encoded input, which is intended to represent card images, followed by the resultant output.

Example 1: Initial Coding

Example 1 shows the coding required at the outset of a document. The encoded input produced the text for the title page, preface, and introduction which follows. It also specifies the alter codes that normally appear preceding the input text. Note that certain alter codes (e.g., DWIDTH, DEPTH, JUST) have been omitted. These specifications, which are in default, are obtained by the program from the operating mode. If the default conditions had not been acceptable for this application, the additional alter codes would have been specified.

```
+PAGE1+
+TITLE+@PL/ZI (F) C@@OMPILER *PROGRAMMER'S *GUIDE
+SUBTITLE+@OM 5502 NM 5521 PT 5523 BAC19@@
+DATE+2 *JUNE 1966
+LIST(1,2,3,4)+*TABLE OF *CONTENTS
+LIST(F)+*LIST OF *FIGURES
+NEWLEVEL+
-S1-*FILE *NO./I/I*S360/S
-I0-*FORM *C28/S6594/S0
-S4-@IBM S@@YSTEM/Z360 *OPERATING *SYSTEM
-I0-@PL/ZI (F)
-I0-P@@ROGRAMMER'S *GUIDE
-I0-360*S/S*N*L/S511
-S4I3-*THE @PL/ZI (F) C@@OMPILER PROVIDES FAST TRANSLATION OF @PL/ZI@@ SOURCE PR
OGRAMS INTO EFFICIENT OBJECT PROGRAMS IN *SYSTEM/Z360 MACHINE LANGUAGE. *THE COM
PILER FUNCTIONS WITHIN THE *OPERATING *SYSTEM/Z360 AND MAY BE USED ON SYSTEMS WH
ERE AT LEAST 45,056 BYTES (44*K) OF CORE STORAGE ARE AVAILABLE FOR THE COMPILER
ITSELF. -P-*THE (*F) *COMPILER ACCEPTS, AS INPUT, *P*L/Z*I SOURCE PROGRAMS FROM
PUNCHED CARDS, PAPER TAPE, MAGNETIC TAPE, OR DIRECT/SACCESS STORAGE DEVICES. */
IT ALSO PROVIDES THE USER WITH POWERFUL AND COMPREHENSIVE DIAGNOSTIC FACILITIES.
-P-*THIS PUBLICATION SERVES TWO PURPOSES. *FIRST, IT SHOULD BE USED IN CONJUNCTI
ON WITH, AND AS A SUPPLEMENT TO, THE PUBLICATION @$IBM S@@YSTEM/Z360  *OPERATING
 *SYSTEM, @PL/ZI: L@@ANGUAGE *SPECIFICATIONS$$, *FORM *C28/S6571, WHICH PROVIDES
 A COMPLETE DESCRIPTION OF *P*L/Z*I. *THE PRESENT MANUAL DESCRIBES CERTAIN RESTR
ICTIONS AND CONVENTIONS THAT APPLY TO PROGRAMS TO BE USED WITH THE (*F) *COMPILE
R.  *SECOND, THIS MANUAL PRESENTS ADDITIONAL INFORMATION OF INTEREST TO USERS OF
 THE COMPILER, SUCH AS OPTIONS AVAILABLE TO THE USER, AND CERTAIN RELATIONSHIPS
BETWEEN THE COMPILER AND THE *OPERATING *SYSTEM. -H1-@PREFACE-P-T@@HIS PUBLICATI
ON DESCRIBES THE FACILITIES PROVIDED BY THE @PL/ZI (F) C@@OMPILER, WHICH FUNCTIO
NS UNDER *OPERATING *SYSTEM/Z360.  *IT COVERS THOSE TEMPORARY LANGUAGE RESTRICTI
ONS AND OPERATING CONSIDERATIONS THAT THE USER MUST TAKE INTO ACCOUNT IN USING T
HE (*F) *COMPILER.  *THE READER IS ASSUMED TO HAVE A WORKING KNOWLEDGE OF BOTH *
P*L/Z*I AND *OPERATING *SYSTEM/Z360, AND HE SHOULD THEREFORE BE FAMILIAR WITH TH
E FOLLOWING PUBLICATIONS: +DEFINE?1=S1I3J3+-?1-$@IBM S@@YSTEM/Z360 *OPERATING *S
YSTEM, @PL/ZI: L@@ANGUAGE *SPECIFICATIONS$$, *FORM *C28/S6571-?1-$@IBM S@@YSTEM/
Z360 *OPERATING *SYSTEM, *CONCEPTS AND *FACILITIES$$, *FORM *C28/S6535-?1-$@IBM
S@@YSTEM/Z360 *OPERATING *SYSTEM, *JOB *CONTROL *LANGUAGE$$, *FORM *C28/S6539-?1
-$@IBM S@@YSTEM/Z360 *OPERATING *SYSTEM, *OPERATING *CONSIDERATIONS$$, *FORM *C2
8/S6540-?1-$@IBM S@@YSTEM/Z360 *OPERATING *SYSTEM, *CONTROL *PROGRAM *SERVICES$$
, *FORM *C28/Z6541-?1-$@IBM S@@YSTEM/Z360 *OPERATING *SYSTEM, *SYSTEM *PROGRAMME
R'S *GUIDE$$, *FORM *C28/S6550-?1-$@IBM S@@YSTEM/Z360 *OPERATING *SYSTEM, *STORA
GE *ESTIMATES$$, *FORM *C28/S6551-?1-$@IBM S@@YSTEM/Z360 *OPERATING *SYSTEM, *SY
STEM *GENERATION$$, *FORM *C28/S6554. -H1-@INTRODUCTION@@-H2-@PL/ZI (F) C@@OMPIL
ER -P-*THE @PL/ZI (F) C@@OMPILER PROVIDES FAST TRANSLATION ...USE)
```

continued



PL/I (F) Compiler Programmer's Guide

2 June 1966

OM 5502 NM 5521 PT 5523 BAC19

File No. S360-
Form C28-6594-0

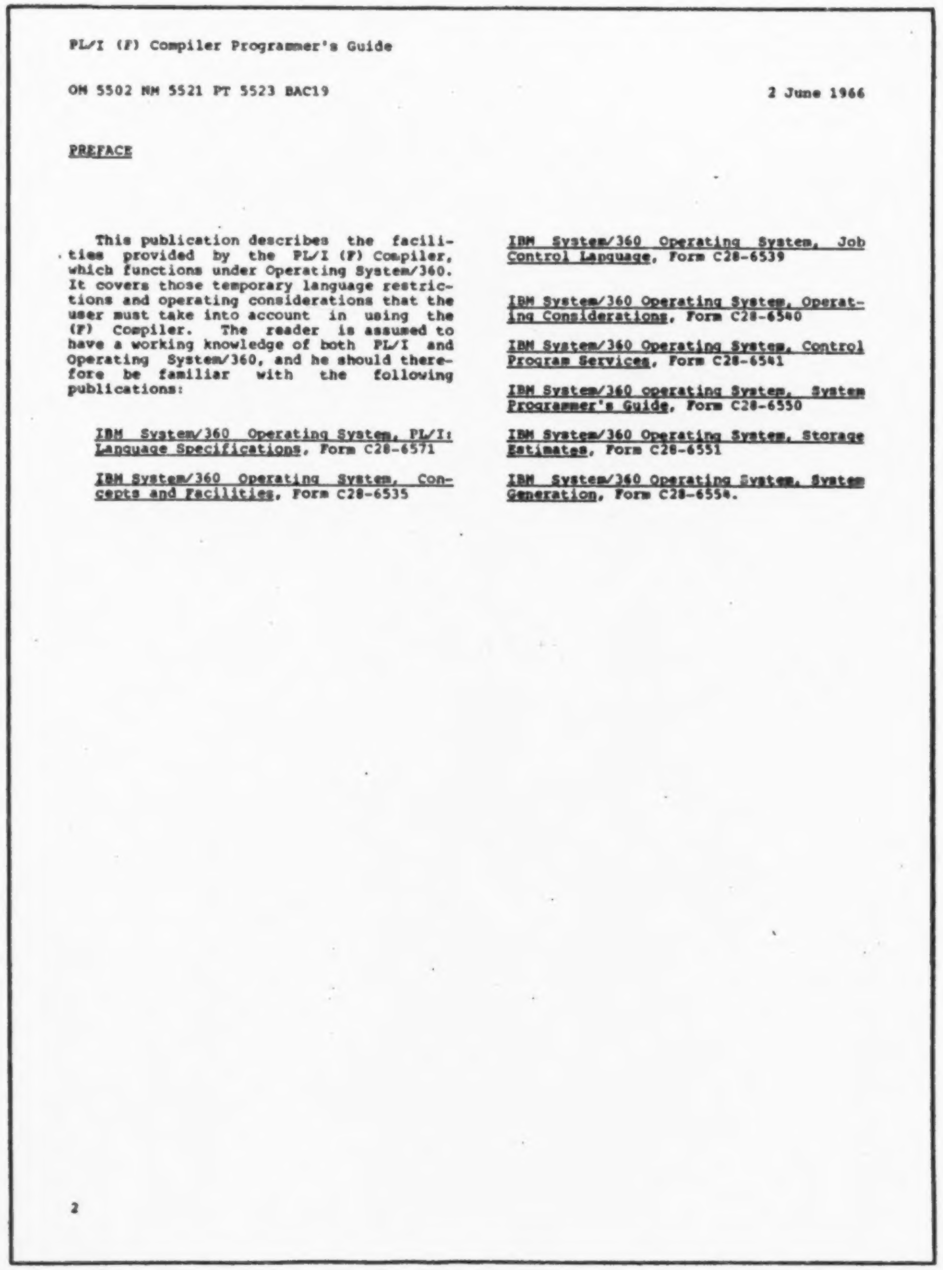
IBM System/360 Operating System
PL/I (F)
Programmer's Guide
360S-NL-511

The PL/I (F) Compiler provides fast translation of PL/I source programs into efficient object programs in System/360 machine language. The compiler functions within the Operating System/360 and may be used on systems where at least 45,056 bytes (44K) of core storage are available for the compilation itself.

The (F) Compiler accepts, as input, PL/I source programs from punched cards, paper tape, magnetic tape, or direct-access storage devices. It also provides the user with powerful and comprehensive diagnostic facilities.

This publication serves two purposes. First, it should be used in conjunction with, and as a supplement to, the publication IBM System/360 Operating System, PL/I: Language Specifications, Form C28-6571, which provides a complete description of PL/I. The present manual describes certain restrictions and conventions that apply to programs to be used with the (F) Compiler. Second, this manual presents additional information of interest to users of the compiler, such as options available to the user, and certain relationships between the compiler and the Operating System.

1



PL/I (F) Compiler Programmer's Guide

OM 5502 NM 5521 PT 5523 BAC19

2 June 1966

PREFACE

This publication describes the facilities provided by the PL/I (F) Compiler, which functions under Operating System/360. It covers those temporary language restrictions and operating considerations that the user must take into account in using the (F) Compiler. The reader is assumed to have a working knowledge of both PL/I and Operating System/360, and he should therefore be familiar with the following publications:

IBM System/360 Operating System, PL/I: Language Specifications, Form C28-6571

IBM System/360 Operating System, Concepts and Facilities, Form C28-6535

IBM System/360 Operating System, Job Control Language, Form C28-6539

IBM System/360 Operating System, Operating Considerations, Form C28-6540

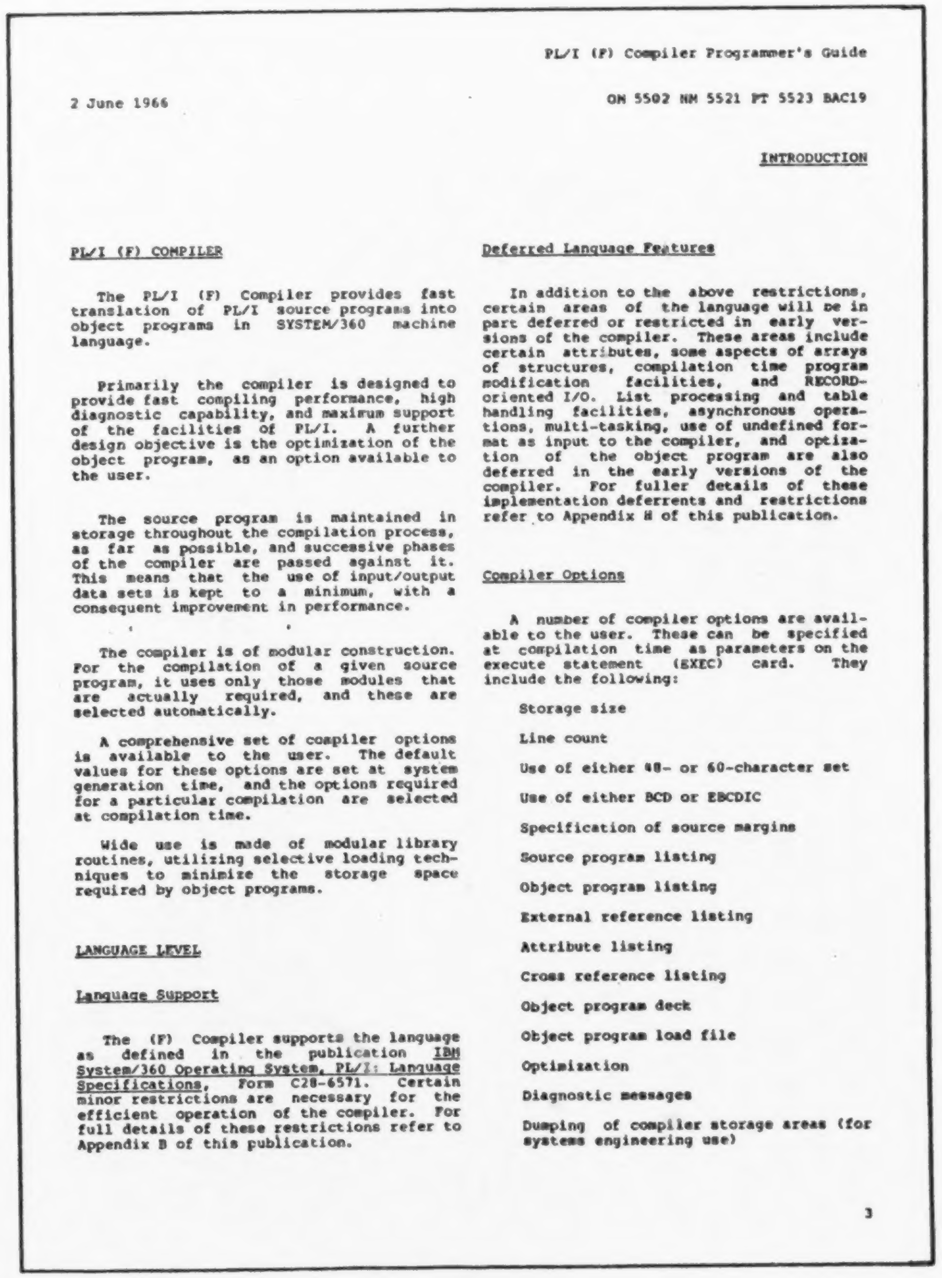
IBM System/360 Operating System, Control Program Services, Form C28-6541

IBM System/360 operating System, System Programmer's Guide, Form C28-6550

IBM System/360 Operating System, Storage Estimates, Form C28-6551

IBM System/360 Operating System, System Generation, Form C28-6554.

2



PL/I (F) Compiler Programmer's Guide

2 June 1966

ON 5502 NM 5521 PT 5523 BAC19

INTRODUCTION

PL/I (F) COMPILER

The PL/I (F) Compiler provides fast translation of PL/I source programs into object programs in SYSTEM/360 machine language.

Primarily the compiler is designed to provide fast compiling performance, high diagnostic capability, and maximum support of the facilities of PL/I. A further design objective is the optimization of the object program, as an option available to the user.

The source program is maintained in storage throughout the compilation process, as far as possible, and successive phases of the compiler are passed against it. This means that the use of input/output data sets is kept to a minimum, with a consequent improvement in performance.

The compiler is of modular construction. For the compilation of a given source program, it uses only those modules that are actually required, and these are selected automatically.

A comprehensive set of compiler options is available to the user. The default values for these options are set at system generation time, and the options required for a particular compilation are selected at compilation time.

Wide use is made of modular library routines, utilizing selective loading techniques to minimize the storage space required by object programs.

LANGUAGE LEVEL

Language Support

The (F) Compiler supports the language as defined in the publication IBM System/360 Operating System, PL/I: Language Specifications, Form C28-6571. Certain minor restrictions are necessary for the efficient operation of the compiler. For full details of these restrictions refer to Appendix B of this publication.

Deferred Language Features

In addition to the above restrictions, certain areas of the language will be in part deferred or restricted in early versions of the compiler. These areas include certain attributes, some aspects of arrays of structures, compilation time program modification facilities, and RECORD-oriented I/O. List processing and table handling facilities, asynchronous operations, multi-tasking, use of undefined format as input to the compiler, and optization of the object program are also deferred in the early versions of the compiler. For fuller details of these implementation deferrents and restrictions refer to Appendix H of this publication.

Compiler Options

A number of compiler options are available to the user. These can be specified at compilation time as parameters on the execute statement (EXEC) card. They include the following:

Storage size

Line count

Use of either 48- or 60-character set

Use of either BCD or EBCDIC

Specification of source margins

Source program listing

Object program listing

External reference listing

Attribute listing

Cross reference listing

Object program deck

Object program load file

Optimization

Diagnostic messages

Dumping of compiler storage areas (for systems engineering use)

continued

44

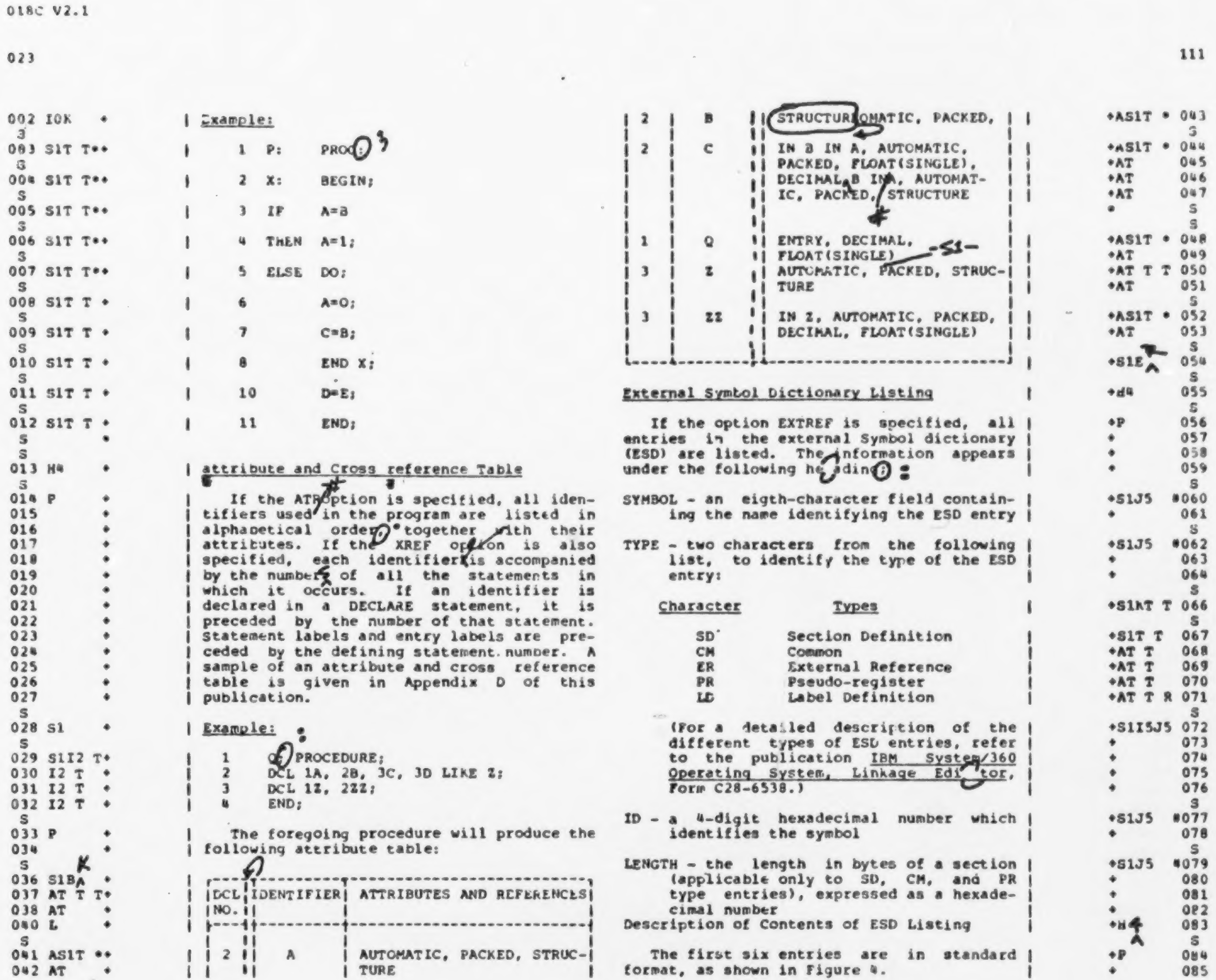
Example 2: Updating Text

Example 2 shows the updating procedure. To clarify the example, the initial input is included. The text page and user reference listing it created follow the initial input. Note how the user reference listing is used to update the lines that are omitted or partially show on the text page. Editorial markings indicate the errors. The marked-up copy is followed by the update codes and the corrected output.

```
+PAGE18+ /
+SETTAB1=5,2=8,3=14+ /
-IOK-$*EXAMPLE:$$ /
-S1T1-1-T-*P:-T-@PROC:@@ /
-S1T1-2-T-*X:-T-@BEGIN; /
-S1T1-3-T-IF-T-A=B /
-S1T1-4-T-THEN-T-A=1; /
-S1T1-5-T-ELSE-T-DO; /
-S1T1-6-T3-A=O; /
-S1T1-7-T3-C=B; /
-S1T1-8-T3-END X; /
-S1T1-10-T3-D=E; /
-S1T1-11-T3-END;@@ /
-H4-ATTRIBUTE AND *CROSS REFERENCE *TABLE /
-P-*IF THE @ATR@@OPTION IS SPECIFIED, ALL IDENTIFIERS USED IN THE PROGRAM ARE  L
ISTED IN ALPHABETICAL ORDER, TOGETHER WITH THEIR ATTRIBUTES. *IF THE @XREF@@  OP
TION IS ALSO SPECIFIED, EACH IDENTIFIERRIS ACCOMPANIED BY THE NUMBER OF ALL   TH
E STATEMENTS IN WHICH IT OCCURS. *IF AN IDENTIFIER IS DECLARED IN A @DECLARE@@
STATEMENT, IT IS PRECEDED BY THE NUMBER OF THAT STATEMENT. *STATEMENT LABELS  AN
D ENTRY LABELS ARE PRECEDED BY THE DEFINING STATEMENT NUMBER. *A SAMPLE OF AN  A
TTRIBUTE AND CROSS REFERENCE TABLE IS GIVEN IN *APPENDIX *D OF THIS PUBLICATION.
-S1-$*EXAMPLE:$$
-S1I2-1-T2-*Q; @PROCEDURE;/
-I2-2-T2-DCL 1A, 2B, 3C, 3D LIKE Z;/
-I2-3-T2-DCL 1Z, 2ZZ;
-I2-4-T2-END;@@/
-P-*THE FOREGOING PROCEDURE WILL PRODUCE THE FOLLOWING ATTRIBUTE TABLE:/
+SETTAB1=2,2=6,3=18+ /
-S1B1,6,16,43+ /
-AT1-@DCL-T-IDENTIFIER-T-ATTRIBUTES AND REFERENCES /
-AT1-NO.@@ /
+SETTAB1=3,2=10,3=18+ /
-L- /
-AS1T1-2-T-@A-T-AUTOMATIC, PACKED, STRUC/S /
-AT3-TURE /
-AS1T1-2-T-B-T-IN A, AUTOMATIC, PACKED, /
-AT3-STRUCTURE /
-AS1T1-2-T-C-T-IN B IN A, AUTOMATIC,/
-AT3-PACKED, FLOAT(SINGLE),/
-AT3-DECIMAL B INA, AUTOMAT/S/
-AT3-IC, PACKED, STRUCTURE /
-AS1T1-1-T-Q-T-ENTRY, DECIMAL, /
-AT3-FLOAT(SINGLE)/
-AT1-3-T-Z-T-AUTOMATIC, PACKED, STRUC/S/
-AT3-TURE /
-AS1T1-3-T-ZZ-T-IN Z, AUTOMATIC, PACKED, /
-AT3-DECIMAL, FLOAT(SINGLE)@@ /
-S1E- /
-H4-*EXTERNAL *SYMBOL *DICTIONARY *LISTING/
-p-*IF THE OPTION @EXTREF@@ IS SPECIFIED, ALL ENTRIES IN THE EXTERNAL *SYMBOL*
DICTIONARY (*E*S*D*) ARE LISTED. *THE INFORMATION APPEARS UNDER THE FOLLOWING HE
ADING;/
-S1J5-@SYMBOL@@/I/S/IAN EIGTH/SCHARACTER FIELD CONTAINING THE NAME IDENTIFYING
THE *E*S*D ENTRY/
-S1J5-@TYPE@@/I/S/ITWO CHARACTERS FROM THE FOLLOWING LIST, TO IDENTIFY THE TYPE
OF THE *E*S*D ENTRY:/
+SETTAB1=5,2=9,3=19,4=24+ /
-S1KT1-$*CHARACTER$$-T4-$*TYPES$$/
-S1T2-@SD-T-S@@ECTION *DEFINITION/
-AT2-*C*M-T-*COMMON/
-AT2-*E*R-T-*EXTERNAL *REFERENCE/
-AT2-*P*R-T-*PSEUDO/SREGISTER/
-AT2-*L*D-T-*LABEL *DEFINITION-R- /
```

continued

```
-S1I5J5-(*FOR A DETAILED DESCRIPTION OF THE DIFFERENT TYPES OF *E*S*D ENTRIES,
REFER TO THE PUBLICATION $@IBM S@@YSTEM/Z360 *OPERATING *SYSTEM, *LINKAGE *EDI
TOR$$, *FORM *C28/S6538.)/
-S1J5-*I*D/I/S/IA 4/SDIGIT HEXADECIMAL NUMBER WHICH IDENTIFIES THE SYMBOL/
-S1J5-@LENGTH@@/I/S/ITHE LENGTH IN BYTES OF A SECTION (APPLICABLE ONLY TO *S*D,
*C*M, AND *P*R TYPE ENTRIES), EXPRESSED AS A HEXADECIMAL NUMBER /
-H-*DESCRIPTION OF *CONTENTS OF @ESD L@@ISTING /
-P-*THE FIRST SIX ENTRIES ARE IN STANDARD FORMAT, AS SHOWN IN *FIGURE 4./
```



```
018C V2.1

023                                                                          111

002 IOK    *    | Example:
 3
003 S1T T**     |     1  P:    PROC;
 S
004 S1T T**     |     2  X:    BEGIN;
 S
005 S1T T**     |     3  IF    A=B
 S
006 S1T T**     |     4  THEN  A=1;
 S
007 S1T T**     |     5  ELSE  DO;
 S
008 S1T T *     |     6        A=0;
 S
009 S1T T *     |     7        C=B;
 S
010 S1T T *     |     8        END X;
 S
011 S1T T *     |     10       D=E;
 S
012 S1T T *     |     11       END;
 S         *
 S
013 H4     *    | attribute and Cross reference Table
 S
014 P      *    |    If the ATR option is specified, all iden-
015        *    | tifiers used in the program are  listed  in
016        *    | alphabetical  order, together  with  their
017        *    | attributes.  If the  XREF  option  is  also
018        *    | specified,  each  identifier is accompanied
019        *    | by the numbers of  all  the  statements  in
020        *    | which  it  occurs.   If  an  identifier  is
021        *    | declared  in  a  DECLARE  statement,  it  is
022        *    | preceded  by  the number of that statement.
023        *    | Statement labels and entry labels are  pre-
024        *    | ceded  by the defining statement number.  A
025        *    | sample of an attribute and cross  reference
026        *    | table  is  given  in  Appendix  D  of  this
027        *    | publication.
 S
028 S1     *    | Example:
 S
029 S1I2 T*     |   1     Q: PROCEDURE;
030 I2 T   *    |   2     DCL 1A, 2B, 3C, 3D LIKE Z;
031 I2 T   *    |   3     DCL 1Z, 2ZZ;
032 I2 T   *    |   4     END;
 S
033 P      *    |    The foregoing procedure will produce the
034        *    | following attribute table:
 S
036 S1B    *    | |---||----------|-------------------------|
037 AT T T*     | |DCL||IDENTIFIER| ATTRIBUTES AND REFERENCES|
038 AT     *    | |NO.||          |                          |
040 L      *    | |---||----------|--------------------------|
 S              | |   ||          |                          |
041 AS1T **     | | 2 ||    A     | AUTOMATIC, PACKED, STRUC-|
042 AT     *    | |   ||          | TURE                     |
```

```
| 2 |    B    || STRUCTURE AUTOMATIC, PACKED,  | |        *AS1T * 043
|   |         ||                               |                    S
| 2 |    C    || IN B IN A, AUTOMATIC,         | |        *AS1T * 044
|   |         || PACKED, FLOAT(SINGLE),        | |        *AT     045
|   |         || DECIMAL, B IN A, AUTOMAT-     | |        *AT     046
|   |         || IC, PACKED, STRUCTURE         | |        *AT     047
|   |         ||                               |          *        S
|   |         ||                               |                   S
| 1 |    Q    || ENTRY, DECIMAL,               | |        *AS1T * 048
|   |         || FLOAT(SINGLE)                 | |        *AT     049
| 3 |    Z    || AUTOMATIC, PACKED, STRUC-     | |        *AT T T 050
|   |         || TURE                          | |        *AT     051
|   |         ||                               |                   S
| 3 |   ZZ    || IN Z, AUTOMATIC, PACKED,      | |        *AS1T * 052
|   |         || DECIMAL, FLOAT(SINGLE)        | |        *AT     053
|   |         ||                               |                   S
|---|---------||-------------------------------| |        *S1E    054
                                                                   S
External Symbol Dictionary Listing               |        *H4     055
                                                                   S
   If the option EXTREF is  specified,  all      |        *P      056
entries  in  the external Symbol dictionary      |        *       057
(ESD) are listed.  The information  appears      |        *       058
under the following headings:                    |        *       059
                                                                   S
SYMBOL - an  eigth-character field contain-      |        *S1J5  *060
     ing the name identifying the ESD entry      |        *       061
                                                                   S
TYPE - two characters  from  the  following      |        *S1J5  *062
     list,  to identify the type of the ESD      |        *       063
     entry:                                      |        *       064
                                                                   S
   Character          Types                      |        *S1KT T 066
                                                                   S
       SD        Section Definition              |        *S1T T  067
       CM        Common                          |        *AT T   068
       ER        External Reference              |        *AT T   069
       PR        Pseudo-register                 |        *AT T   070
       LD        Label Definition                |        *AT T R 071
                                                                   S
     (For a  detailed  description  of  the      |        *S1I5J5 072
     different  types of ESD entries, refer      |        *       073
     to  the  publication  IBM  System/360       |        *       074
     Operating  System,  Linkage  Editor,        |        *       075
     Form C28-6538.)                             |        *       076
                                                                   S
ID - a  4-digit  hexadecimal  number  which      |        *S1J5  *077
     identifies the symbol                       |        *       078
                                                                   S
LENGTH - the  length  in bytes of a section      |        *S1J5  *079
     (applicable only to  SD,  CM,  and  PR      |        *       080
     type  entries), expressed as a hexade-      |        *       081
     cimal number                                |        *       082
Description of Contents of ESD Listing           |        *H4     083
                                                                   S
   The first six entries  are  in  standard      |        *P      084
format, as shown in Figure 4.                    |        *       085
```

continued

```
                   ALTER CODES, EXPANDED EDIT CODES, FOOTNOTES, AND ERROR MESSAGES

PAG LIN                CODE

018 001 *SETTAB1=5,2=8,3=14*
018 003 -S1T1-1-T-P:-T-PROC:
018 004 -S1T1-2-T-X:-T-BEGIN;
018 005 -S1T1-3-T-IF-T-A=B
018 006 -S1T1-4-T-THEN-T-A=1;
018 007 -S1T1-5-T-ELSE-T-DO;
018 008 -S1T1-6-T3-A=O;
018 009 -S1T1-7-T3-C=B;
018 010 -S1T1-8-T3-END X;
018 011 -S1T1-10-T3-D=E;
018 012 -S1T1-11-T3-END;
018 029 -S1I2-1-T2-Q: PROCEDURE;
018 030 -I2-2-T2-DCL 1A, 2B, 3C, 3D LIKE Z;
018 031 -I2-3-T2-DCL 1Z, 2ZZ;
018 032 -I2-4-T2-END;
018 035 *SETTAB1=2,2=6,3=18*
018 036 -S1B1,6,16,43*
018 037 -AT1-DCL-T-IDENTIFIER-T-ATTRIBUTES AND REFERENCES
018 038 -AT1-NO.
018 039 *SETTAB1=3,2=10,3=18*
018 041 -AS1T1-2-T-A-T-AUTOMATIC, PACKED, STRUC-
018 042 -AT3-TURE
018 043 -AS1T1-2-T-B-T-IN A, AUTOMATIC, PACKED, -T3-STRUCTURE
018 044 -AS1T1-2-T-C-T-IN B IN A, AUTOMATIC,
018 045 -AT3-PACKED, FLOAT(SINGLE),
018 046 -AT3-DECIMAL, B IN A, AUTOMAT-
018 047 -AT3-IC, PACKED, STRUCTURE
018 048 -AS1T1-1-T-Q-T-ENTRY, DECIMAL,
018 049 -AT3-FLOAT(SINGLE)
018 050 -AT1-3-T-Z-T-AUTOMATIC, PACKED, STRUC-
018 051 -AT3-TURE
018 052 -AS1T1-3-T-ZZ-T-IN Z, AUTOMATIC, PACKED,
018 053 -AT3-DECIMAL, FLOAT(SINGLE)
018 065 *SETTAB1=5,2=9,3=17,4=24*
018 066 -S1T1-Character-T4-Items
018 067 -S1T2-SD-T-Section Definition
018 068 -AT2-CM-T-Common
018 069 -AT2-ER-T-External Reference
018 070 -AT2-PR-T-Pseudo-register
018 071 -AT2-LD-T-Label Definition--
<*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*>

        T1 0.102  n NOT SPECIFIED FOR -nn- CODE; -I0- ASSUMED
018 083 -n-Description of Contents of ESD Listing
<*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*><*>
```

-AS1T1-2-T-*D-T-*I*N

A

K

#

4

```
+18003R*C:+*C;/
+13RA+*A/
+13R R+ *R/
+14R*R+*R /
+16R.+,/
+18RRR+R /
+19RR+RS/
+29R;+:/
+36R1+1*K/
+36R6+5/
+43R-T3-+-AT3-/
+46R*L+*L -AS1T1-2-T-*D-T-*I*N /
+46R*N+*N /
+50R*A+*A*S1/
+54R*E+*E*R/
+59RE A+EA/
+59R;+:/
+75RI +I/
+85R-H-+-H4-/
```

continued

Example:

```
     1  P:     PROC;
     2  X:     BEGIN;
     3  IF     A=B
     4  THEN   A=1;
     5  ELSE   DO;
     6         A=O;
     7         C=B;
     8         END X;
     10        D=E;
     11        END;
```

Attribute and Cross Reference Table

If the ATR option is specified, all identifiers used in the program are listed in alphabetical order, together with their attributes. If the XREF option is also specified, each identifier is accompanied by the numbers of all the statements in which it occurs. If an identifier is declared in a DECLARE statement, it is preceded by the number of that statement. Statement labels and entry labels are preceded by the defining statement number. A sample of an attribute and cross reference table is given in Appendix D of this publication.

Example:

```
  1    Q: PROCEDURE;
  2    DCL 1A, 2B, 3C, 3D LIKE Z;
  3    DCL 1Z, 2ZZ;
  4    END;
```

The foregoing procedure will produce the following attribute table:

DCL NO.	IDENTIFIER	ATTRIBUTES AND REFERENCES
2	A	AUTOMATIC, PACKED, STRUCTURE
2	B	IN A, AUTOMATIC, PACKED, STRUCTURE
2	C	IN B IN A, AUTOMATIC, PACKED, FLOAT(SINGLE), DECIMAL
2	D	IN B IN A, AUTOMATIC, PACKED, STRUCTURE
1	Q	ENTRY, DECIMAL, FLOAT(SINGLE)
3	Z	AUTOMATIC, PACKED, STRUCTURE
3	ZZ	IN Z, AUTOMATIC, PACKED, DECIMAL, FLOAT(SINGLE)

External Symbol Dictionary Listing

If the option EXTREF is specified, all entries in the external Symbol dictionary (ESD) are listed. The information appears under the following heading:

SYMBOL - an eigth-character field containing the name identifying the ESD entry

TYPE - two characters from the following list, to identify the type of the ESD entry:

Character	Types
SD	Section Definition
CM	Common
ER	External Reference
PR	Pseudo-register
LD	Label Definition

(For a detailed description of the different types of ESD entries, refer to the publication IBM System/360 Operating System, Linkage Editor, Form C28-6538.)

ID - a 4-digit hexadecimal number which identifies the symbol

LENGTH - the length in bytes of a section (applicable only to SD, CM, and PR type entries), expressed as a hexadecimal number

continued

Example 3: Table Coding

```
-H4-*DATA *ELEMENT *DESCRIPTOR (@DED)-P-T@@HE FORMAT OF THE
 DATA ELEMENT DESCRIPTOR (@DED)(F@@IGURE 21) IS AS FOLLOWS:
+SETTAB1=3,2=15,3=32,4=34,5=37,6=40,7=42,8=44,9=49,10=52,11=54,12=58+
-DS1B1,13,30,66-
-AT5-@DED F@@ORMATS (IN *BYTES)
-AT1-*DATA *TYPE-T-*REPRESENTATION-G30,66V38,42,46,52,56-
-AT4-1-T6-2-T8-3-T-4-T11-5-T12-6.....
-AL-
-AT2-*FIXED/SPOINT
-S1AT1-*ARITHMETIC-T-*FLOATING/SPOINT -T-*FLAGS -T6-*P-T8-*Q
-S1AT2-*PACKED/SDECIMAL
-AG13,66-
-AT2-*NUMERIC FIELD -T-*FLAGS-T6-@P-T8-Q-T-W-T11-L-T12-P@@ICTURE
-AL-
-AT2-*UNPICTURED-T-*FLAGS
-AT1-*STRING -G13,66Q42,52,56
-AT2-*PICTURED -T-*FLAGS-T7-*L-T10-*PICTURE
-AEQ-
-S0XF- *FIGURE 21.  *DATA *ELEMENT *DESCRIPTOR (*D*E*D)-R-
-S1- *FLAGS: *AN EIGHT/SBIT ENCODED FORM OF DECLARED INFORMATION (SEE *FIGURE 22
).∕
+SETTAB1=1,2=7,3=9,4=16,5=20,6=23,7=31,8=33,9=39,10=43,11=50,12=53,13=59,14=61,1
5=67,16=68,17=70,18=77+∕
-DS1B5,14,18,29,37,48,57,65,75-
-AT3-0-T-1-T6-2-T8-3-T10-4-T12-5-T14-6-T17-7
-AL-
-AT4-0 -T-*PACKED -T7-*FIXED -T9-*PICTURED -T11-*BIT-T14-0-T17-0-T-= 0
-AT1-0 =-T-*STRING-G14,75-
-AT4-0-T-*ALIGNED -T7-*VARY/S -T9-*UN/S-T11-*CHARAC/S-T14-0-T17-0-T-= 1
-AT7-ING -T9-PICTURED -T11-TER
-AL-
-AT4-0-T-*NON/S -T7-*SHORT -T9-*NUMERIC -T11-*DECI/S -T13-*FIXED -T16-*REAL
-T18-= 0∕
-AT1-1 =-T-*ARITH/S-T5-*STERLING-T9-*FIELD-T11-MAL
-AT2-METIC-G14,75-
-AT4-0-T-*STERLING-T7-*LONG-T9-*CODED-T11-*BINARY -T13-*FLOAT-T15-*COMPLEX-T18-=
 1
-AE-
-AXFT1-*FIGURE 22. *EIGHT/S*BIT *ENCODED *FORM OF *DECLARED *INFORMATION IN *FLA
GS-R- -P- *THE *P BYTE IS THE DECLARED OR DEFAULT PRECISION OF THE DATUM. *MAXIM
UM VALUES ARE:-S1I4-*BINARY/I/I*FIXED: 31-I4-*BINARY/I/I*FLOAT: 53          -I-*D
ECIMAL *FIXED: 15-I-*DECIMAL *FLOAT: 16-P-*THE *Q BYTE IS THE DECLARED OR DEFAUL
T SCALE FACTOR OF THE DATUM IN EXCESS 128 NOTATION (I.E., IF THE IMPLIED FRACTIO
NAL POINT IS BETWEEN THE LAST AND NEXT TO LAST DIGIT, *Q WILL HAVE A VALUE 129).
 *FOR NUMERIC FIELDS, *Q IS THE RESULTANT SCALE FACTOR DERIVED FROM THE APPARENT
 PRECISION...
```

continued

Data Type	Representation	DED Formats (in Bytes)					
		1	2	3	4	5	6.....
Arithmetic	Fixed-point Floating-point Packed-decimal	Flags	P	Q			
	Numeric field	Flags	P	Q	W	L	Picture
String	Unpictured	Flags					
	Pictured	Flags	L		Picture		

Figure 21. Data Element Descriptor (DED)

	0	1	2	3	4	5	6	7	
0 =	String	0	Packed	Fixed	Pictured	Bit	0	0	= 0
		0	Aligned	Varying	Unpictured	Character	0	0	= 1
1 =	Arithmetic	0	Non-Sterling	Short	Numeric Field	Decimal	Fixed	Real	= 0
		0	Sterling	Long	Coded	Binary	Float	Complex	= 1

Figure 22. Eight-Bit Encoded Form of Declared Information in Flags

Data Element Descriptor(DED)

The format of the data element descriptor (DED) (Figure 21) is as follows:

Flags: An eight-bit encoded form of declared information (see Figure 22).

The P byte is the declared or default precision of the datum. Maximum values are:

Binary Fixed: 31
Binary Float: 53
Decimal Fixed: 15
Decimal Float: 16

The Q byte is the declared or default scale factor of the datum in excess 128 notation (i.e., if the implied fractional point is between the last and next to last digit, Q will have a value 129). For numeric fields, Q is the resultant scale factor derived from the apparent precision as specified in the picture, i.e., the number of digit positions after a V picture item as modified by an F (scale factor) item.

The W byte specifies the number of storage units allocated for a numeric field (byte count if decimal, bit count if binary).

The L byte specifies the number of bytes allocated for the picture associated with a numeric field; if the datum is string, L occupies two bytes; if arithmetic, one byte.

The picture specification field contains the picture declared for the datum; if the datum is string, the picture may occupy 1 through 32,767 bytes; if arithmetic, 1 through 255 bytes. If the original picture specification contained iteration factors, it will have been expanded in full.

Label Data

Data of type LABEL takes on the value of statement labels. Label variables occupy a two-word field aligned to a full-word address. The field is used a follows:
end of test

COST ESTIMATE

For a 180 page manual which utilized all but the spelling check features of the system, the computer costs of updating the master file with numerous changes and producing camera-ready copy were $2.50 per page per update.

Charge to user = computer costs + network overhead

CONTENTS—TEXT360

pages

1– 2	Identification & Abstract
3	User Instructions
5–13	I/O
15	Cost—Contents

DESCRIPTIVE TITLE	Synagraphic Mapping Program
CALLING NAME	SYMAP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Harvard University Graduate School of Design
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not distribution
CONTACT	Dr. Charles Drebes, Manager, Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

SYMAP is a computer program designed to allow city planners, geographers and others to produce low cost graphic displays of spatial patterns using standard computer line printers, by producing maps which graphically depict spatially disposed quantitative and qualitative information. It is suited to a broad range of applications, and is provided with numerous options to meet widely varying requirements. Raw data of every kind (physical, social, economic, etc.) when given to the computer may be related, manipulated, weighted, and aggregated in any manner desired. By assigning values to the coordinate locations of data points or data zones, one or more of three types of map may be produced, as specified by the user: conformant (choropleth), contour, and proximal. Potential applications are independent of the scale at which one wishes to display data. Studies (at other universities) have included a living cell, land parcels, blocks, tracts, towns, states, and continents. In each case, a common factor was the spatial distribution of a variable and a need to display the patterns associated with this distribution.

REFERENCES

Reference Manual for Synagraphic Computer Mapping—"SYMAP" (Cambridge, Mass.: Harvard Univ. Grad. Sch. of Design, Comp. Graph. Lab.).

USER INSTRUCTIONS

See the reference listed below.

REFERENCES

Reference Manual for Synagraphic Computer Mapping—"SYMAP" (Cambridge, Mass.: Harvard Univ. Grad. Sch. of Design, Comp. Graph. Lab.).

SAMPLE INPUT and SAMPLE OUTPUT

Interested persons should consult the contact person and the reference listed in the User Instructions.

COST ESTIMATE

Interested persons should consult the contact person.

Charge to user = computer costs + network overhead

CONTENTS—SYMAP

pages

1 Identification & Abstract
3 User Instructions
5 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	FURNIVAL Regression Screen
CALLING NAME	FURNIVAL
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Van D. Nelson, Numerical Analysis and Programming Section, Statistical Laboratory, Ames, Ia. 50010
LANGUAGE	FORTRAN II
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

Investigators often want to select from a large group of independent variables a smaller number to be used as predictors in a regression equation. One possible method is to compute regressions on *all* possible combinations of the variables. However, the total number of possible regressions can become very large. The purpose of this program is to reduce the number of regressions computed to those combinations of variables meeting one or more of four constraints.

Variables may be fixed or forced to appear in every regression, reducing the number of regressions to those combinations of variables containing the fixed variables. For example, consider the five variables 1,2,3,4, and 5, in which 1 and 3 are to be fixed. Then only those combinations using both 1 and 3 are computed; for instance, the combinations (1,3), (1,3,2), and (1,3,2,5) will have regressions computed, but the combinations (2,4,5), (1,2,4), and (3,5) will not, because *both* 1 and 3 are not present. Also, the maximum number of variables appearing in any regression may be limited to any number less than the total number of variables.

In addition, variables may be placed in sets such that, if one variable in a set appears, all variables in that set will appear. In other words, either the entire set of variables or none of the var-

continued

iables in the set is present. For example, let the number of variables be 5, and let variables 1 and 2 be in set A, variable 3 be in set B, and variables 4 and 5 be in set C. Then regressions would be computed for (3), (1,2), (1,2,3), and (1,2,4,5) but not for (1), (1,3), (3,4), or (2,3,4). Essentially, the program treats a set of variables as a single variable, and instead of using combinations of variables it computes regressions on all combinations of sets.

Variables or sets of variables may also be placed in groups such that,if one member of the group is present in a regression, no other member of that group will be present. For example, using the same sets A, B, and C above, regressions would be computed for (1), (2), (1,3), (1,4), and (2,3,5). Regressions would not be computed on (1,2), (1,2,4), or (1,3,4,5). Either no member or only one member of a group can be present. The same conditions would hold if 1,2,3,4, and 5 represented sets instead of variables.

Computed for each regression are (1) the coefficient of colinearity (the determinant of the gross-moments matrix for the independent variables, scaled by dividing by the logical product of the elements in its main diagonal), used as a measure of the degree of associated among the independent variables, including the dummy variables or constants, and (2) the coefficient of determination (R square), computed conventionally from the formula $R^2 = 1-(SSR/SSY)$, where SSR is the sum of squared residuals and SSY is the sum of squares about the mean of the dependent variable.

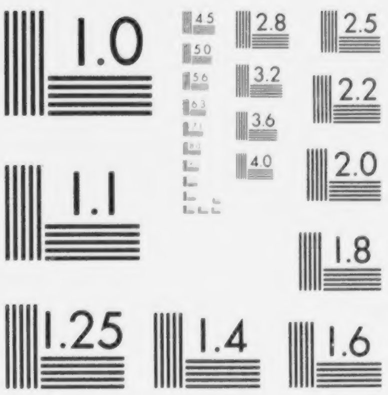
REFERENCES

Nelson, Van D., *User Manual—FURNIVAL Regression Screen*, Statistical Lab., Ames, Ia. (1968). Available from the EIN office for the cost of duplication and mailing.

2 OF 12

ED

064 958


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

USER INSTRUCTIONS

Description of Input

The preparation of the data deck for FURNIVAL is fairly straightforward and completely described by listing the cards as they should appear.

Title Card

This first card must always be present. It consists of any alphanumeric characters in Cols. 1–72. This string of characters will be printed out at the top of the output.

Control Card

This second card also must always be present. If none of the options or constraints of the program is employed, the Control Card is fairly simple

Columns	*Contents*
6–10	number of observations on each variable (right-justified)
19–20	number of independent variable (right-justified)
30	number of dependent variables

Format Card

This is the third and, if necessary, the fourth card; it, too, must always be present. This card is like a regular FORTRAN format statement except the word FORMAT is omitted, and only E- or F-type conversion is permitted. Column 1 contains a left parenthesis followed by the format of the observations, and finally by a right parenthesis. If more than one card is required for the format statement, punch a 1 in Col. 5 of the Control Card.

Observation Card

These cards will follow the Format Card when none of the options or constraints is used. Otherwise, there will be other cards between the Format Card and Observation Cards. There should be a value for each independent variable and each dependent variable, and their placement should correspond with the Format Card. There should be as many observations on each variable as specified in Cols. 6–10 on the Control Card. The independent variables must be listed first, followed by the dependent variables. When constraints are used, the independent variables must be in a certain order. Fixed variables are listed first. If sets are used, those

continued

variables to be placed in the same set must be together. If groups are used, those variables or sets of variables that are to be placed in the same group must be together. If the cards are already punched and the variables are not in the above order, it will be necessary to use the transformation subroutine (see options below).

Done Card

This card must always follow the last card of the Observation Deck, and consists of the four letters DONE, starting in Col. 1. If you want to run more sets of data on the same run, follow the Done Card with the next Title Card and repeat the above process.

Options

The control cards described above must be modified whenever regressions are to be forced through the origin; variables are to be identified with labels supplied by the user; constraints are applied to limit the number of regressions; or the Transformation Subroutine is used. Most of the options require additional information on the Control Card, and some require additional cards.

1. To force regression *through the origin*, punch a 1 in Col. 4 of the Control Card.
2. The program will label independent variables and sets of independent variables as they are encountered, giving them consecutive numbers. It should be noted that each *set* receives only one label regardless of the number of variables in it. Dependent variables are labelled by Y1, Y2, and so on. If these labels are not satisfactory, punch a 1 in Col. 1 of the Control Card and add the following card(s) behind the Format Card. In Col. 1 leave a blank and then use any three characters to label variable 1 (set 1 if you are using sets); continue this process until all independent variables (sets) and all dependent variables are labelled. If this process takes more than 72 columns, start a new card by entering a blank and then three characters. Each label should contain four characters, including the first one as a blank. Any of the other characters may also be a blank; but the first one *must* be.
3. To fix variables, punch the number of fixed variables (sets) in Cols. 39–40 of the Control Card. If you are using sets, the program will print out a limited number of regressions that contain only fixed variables. The first regression contains the first set of fixed variables; the second includes the first two sets of fixed variables, and the kth regression includes the first k sets of fixed variables. The change in R square from one regression to the next will reflect the addition of an en-

continued

tire set of fixed variables. If the effect of adding individual fixed variables is desired, then each fixed variable should be put in its own set. If the above has no value to the user, then all fixed variables should be put in the same set. When fixing variables, care should be taken when counting the number of variables for other places on the Control Card. Columns 19–20 of the Control Card contain the number of independent variables (sets) *not fixed*. Columns 39–40 contain the number of *fixed* variables (sets). Columns 49–50 contain the largest number of independent variables (sets) *not fixed to be included in any regression.*

4. To put independent variables in *sets*, punch a 1 in Col. 2 of the Control Card and add the following card behind the label cards. (If there are no label cards, put it behind the Format Card.) In Cols. 1–2 put the number of variables in set 1; in Cols. 3–4 put the number of variables in set 2; and so on. All independent variables including fixed variables must be put in a set. It is necessary that the variables be in the proper order (see Observation Cards). Notice that, even if only 1 variable is in a set, the number 1 must be punched. If sets are not specified, the program puts each independent variable in its own set.

5. To put independent variables (sets) in *groups*, punch a 1 in Col. 3 of the Control Card and add the following card behind the set card. (If there is no set card, put it behind label cards; if there are no label cards, put it behind the Format Card.) In Cols. 1–2 punch the number of groups that will be formed; in Cols. 3–4 punch the number of variables (sets) in group 1; in Cols. 5–6 punch the number of variables (sets) in group 2; and so on until all independent variables (sets) not fixed have been included in a group. Even if a group contains only one variable (set), the number 1 must be punched. It is important to remember that, in grouping, fixed variables (sets) are ignored and group assignment begins with the first *nonfixed* variable (set).

6. To place a maximum on the number of *(nonfixed)* variables (sets) in any regression, punch the maximum number in Cols. 49–50 of the Control Card.

7. To use the Transformation Subroutine, punch the total number of variables (independent and dependent) before transformations (blank means that they are the same as after transformations) in Cols. 59–60 of the Control Card. For each variable to be transformed, make a card as described below and put them directly in front of the Observation Deck. Count the number of trans-

continued

formations made and punch this number in Cols. 61–62 of the Control Card. The variables are considered to be numbered from 1 to the total number of variables, starting with variables on the left edge of the observation data card. In Cols. 1–2 of a new card put the code number of the transformation desired; in Cols. 3–4 put the number of the resultant (transformed) variable; in Cols. 5–6 put the number of the first variable in the transformation (on the right of the equals sign); in Cols. 7–8 put the number of the second variable (if 2 variables are needed in the transformation); in Cols. 7–14 put the constant term punched with a decimal point. For each transformation punch such a card. See Table II for the types of transformations available and their code numbers. For example, to interchange variables 1 and 6 (because 6 is going to be fixed and must appear as the first independent variable), punch a 2 in Col. 62 of the Control Card for the 2 necessary transformations, then in a new card punch a 1 in Col. 2 for transformation code 01, put a 1 in Col. 4, and a 6 in Col. 6. This makes variable 1 equal to variable 6. Start another card and punch 12 in Cols. 1–2, put 6 in Col. 4, and a 1 in Col. 6. This makes variable 6 equal to the original value of variable 1.

All of the card formats and option cards are summarized in Table I.

TABLE I. Control and Data Cards for FURNIVAL Regression Screen.

Card	*Columns*		*Contents*
1	1–72		[a]job title
2	1	NTF	label card flag (blank: labels furnished by program; 1: labels on card 5)
	2	NRF	set card flag (blank: all sets have one variable; 1: set sizes on card 7)
	3	NGF	group card flag (blank: all groups have one set; 1: group sizes on card 8)
	4	NIF	intercepts flag (blank: intercepts; 1: no intercepts)
	5	NFF	format-card flag (blank: no second data format card; 1: more data format on card 4)
	6–10	[a]NOB	number of observations (>3)

continued

Card	*Columns*		*Contents*
	19–20	^{a}NXS	number (of sets) of independent variables (after transformation) not fixed (<31)
	30	NY	number of dependent variables (<9) (blank: one)
	39–40	NF	number (of sets) of fixed variables (<18) (blank: 0)
	49–50	LM	largest number (of sets) of independent variables not fixed to be included in any regression (<18) (blank: no special restriction)
	59–60	NVR	total number of variables before transformations (<51) (blank: same as after transformation)
	61–62	NTRAN	number of transformations per observation (blank: no transformations)
3	1		^{a}left parenthesis
	2–72		^{a}format of data followed by right parenthesis, if data format is contained on this card only
4	1–72		continuation of data format followed by right parenthesis (omit card 4 if Col. 5 of card 2 is blank)
5	1–4		label for first set of independent variables
	5–8		same for second set
	9–72		labels for other sets of independent variables followed by labels for dependent variable(s)—four columns for each label (omit card 5 if Col. 1 of card 2 is blank)
6	1–72		continuation of labels (omit card 6 unless needed and if Col. 1 of card 2 is blank)
7	1–2		number of variables in first set of independent variables
	3–4		same for second set

continued

Card	*Columns*	*Contents*
	5–72	number of variables for remaining sets of independent variable—two columns for each set (omit card 7 if Col. 2 of card 2 is blank)
8	1–2	number of groups of independent variables
	3–4	number of sets of independent variables in first group
	5–6	number of sets of independent variables in second group
	9–72	number of sets of independent variables for remaining groups—two columns for each group (omit card 8 of Col. 3 if card 2 is blank)
9		transformation control cards (omit these cards if Cols. 61–62 of card 2 are blank or zero). For *each* transformation, there must be *one* transformation control card. The variables are considered to be numbered from 1 to NVR from the left edge of the observation data card. At the end of the transformations the independent variables must occupy positions 1 to (NF+NXS) and the dependent variables occupy positions (NF +NXS+1) to (NF+NXS+LY)
	1–2	number of the transformation, chosen from list
	3–4	number of the resultant (transformed) variable
	5–6	number of the first variable in the transformation (on the right of the = sign)
	7–8	number of the second variable (if two variables are involved in the transformation)
	7–14	constant term (c) punched with decimal point

continued

Card	Columns	Contents
10	1–80	[a]observational data in format specified on card 5 (3 and 4). If more than one data card is needed for each observation, the set of cards for each observation must be together and in the order prescribed by the format. There can be no more than one observation per card. Starting from the left of the card, the *independent* variables are punched *first*, then the *dependent*
11	1–4	[a]DONE. This indicates that all of the data cards have been read.

---

[a]Items which must be included with a nonzero value for every job. Blank space is indicated by "b."

TABLE II. Descriptions of Transformations and Codes

The following transformations may be called by the numbers given on the left (entered in Cols. 1–2 of card 9).

- t resultant variable (Cols. 3–4 of card 9)
- u first variable in the transformation (Cols. 5–6 of card 9)
- v second variable in the transformation, if used (Cols. 7–8 of card 9)
- c constant, punched with the decimal point (Cols. 7–14 of card 9)

Number	*Transformation*
01	$t = u$
02	$t = c * u$
03	$t = c + u$
04	$t = u + v$
05	$t = u * v$
06	$t = u / v$
07	$t = 1 / u$
08	$t = u^c$
09	$t = \ln_c u$
10	$t = c^u$
11	$t = \log_{10} u$

continued

Number	*Transformation*
12	t = original value of u regardless of previous transformations on u
13	t = sin (u)
14	t = cos (u)
15	t = u / c
16	t = |u|
17	call user-written subroutine TROBS

For complex transformations the use of these simple algebraic forms may become unwieldy. The user may code his own transformations as follows.

(a) Call for transformation 17 on a transformation 17 on a transformation control card

(b) Write a FORTRAN II subroutine called TROBS (X), which must start as follows.

```
SUBROUTINE TROBS (X)
DIMENSION X(75)
```

where X(1) to X(NVR) contain a single observation on the variables. This subroutine will be called each time an observation is read in. The user may still call for other observations as well.

Example: This subroutine sums variables 1 through 4 to form variable 5 and forms variable 6 as $(1-x_5{}^2)^{1/2}$

```
SUBROUTINE TROBS (S)
DIMENSION X(75)
X(5) = X(1) + X(2) + X(3) + X(4)
X(6) = SQRT (1.-X(5)**2)
RETURN
END
```

Description of Output

The initial phase of the program reads the information on the Control Card, generates labels and constants not supplied by the user, and makes an advance computation of the number of regressions to be produced. All of this information is then printed out on the first page of output. If any restrictions to the various variables of the program are violated, an asterisk is printed and the error message CONTROL CARD ERROR, CHECK PREVIOUS LISTING FOR ASTERISKS is given. No further computations are made and no other output is printed for a problem in which an error is detected.

When no errors are detected, the program prints out the first three sets of observations. If the transformation subroutine is used, the program also prints out the first transformed sets of variables; hence, the listing may be used to check both the format statement and the transformations made. If data cards are missing, if the number of observations reported on the Control Card is too large, if the format statement causes misreading of the Observation Cards, or if the Done Card is missing, a suitable diagnostic message will be printed; otherwise, the last set of observations is printed and computation of the regression begins.

The major output of the program, the results of the regression computations, forms an array with one line (row) for each regression (see Sample Output). The leading element in each row is the number of sets of independent variables (not fixed) included in each regression. This is followed by the labels or indices for the sets. Then comes the coefficient of colinearity (described later) and, finally, the R square for each dependent variable.

When the fixed constraint has been used, the two types of regressions produced are (1) the auxiliary regressions (containing only fixed variables), and (2) the primary regressions (containing all fixed variables plus one or more sets of variables that are not fixed). The auxiliary regressions are listed first and can be identified by the zeroes shown for number of sets of variables not fixed. Next are the primary regressions in blocks, separated by blank lines, each block composed of regressions containing the same number of sets of independent variables. The labels for fixed variables are included in all primary regressions. The last regression printed includes every independent variable even though this regression may have been screened by the constraints.

For each regression, there follow the coefficients of colinearity and determination. For the colinearity coefficient, a value of 1 would indicate no correlation, and a value of zero would indicate

continued

perfect linear dependence. However, a large negative exponent in the E type of representation is a warning that the matrix is ill conditioned.

REFERENCES

Nelson, Van D., *User Manual—FURNIVAL Regression Screen*, Statistical Lab., Ames, Ia. (1968). Available from the EIN office for the cost of duplication and mailing.

SAMPLE INPUT

```
//STEP1  EKEC PGM=FURN,TIME=(,10),REGION=128K
//FT05F001 DD UNIT=DISK,SPACE=(800,(200,10))
//FT03F001 DD SYSOUT=A,DCB=(RECFM=FBA,LRECL=133,BLKSIZE=3325,BUFNO=2)
//FT01F001 DD *
12 IND. VAR..GROUP1=SET1.GROUP2=SETS2-5,GROUP3=SETS6-7
 11    62            7         1                            13
(3F3.1,F4.1,2F3.1,F4.1,2F3.1,F4.1,F3.1,F3.0,39X,F2.0)
 4 1 1 2 1 1 2
 3 1 4 2
10   1 26  68 43 34  26 50 31  96 35 122                                       51
20   4 20  40 31 15  18 54 10  10 56 314                                       51
10   1 25  62 30 75  30 46 16  26 66 436                                       53
20   4 19  35 42  6  14 54  3   1 55 302                                       54
30   9 18  32 28 36  19 54 15  22 42 176                                       54
20   4 15  22 33 25  13 57  4   2 48 230                                       55
30   9 16  26 34 10  22 51  5   2 85 722                                       55
20   4 22  48 42 39  22 56 30  90 25  62                                       55
20   4 15  22 40 12  21 55  7   5 58 336                                       56
20   4 17  29 16 28  27 49 50 250 28  78                                       56
25   6 17  29 40 47  20 56 15  22 41 168                                       57
15   2 18  32 46 40  20 49 18  32 41 168                                       57
20   4 14  20 24 27  29 47 11  12 46 212                                       57
15   2 13  17  1 25  41 28 36 130 32 102                                       57
20   4 21  44 28 34  20 56 24  58 35 122                                       59
             .         .        .                                               .
             .         .        .                                               .
             .         .        .                                               .

40  16 28  78 30 32  38 36 35 122 39 152                                       68
50  25 34 116 43 51  37 39 36 130 41 168                                       68
20   4 32 102 39 51  14 58 28  78 27  73                                       68
50  25 12  14  1  6  38 34 33 109 33 109                                       69
20   4 11  12 13 23   6 64 11  12 35 122                                       69
30   9 15  22 14 36   8 60  7   5 43 185                                       69
30   9 18  32 14 52  34 43 42 176 30  90                                       70
40  16 17  29  3  8  48 33 49 240 29  84                                       71
20   4 18  32 22 56  18 54 18  32 33 109                                       71
40  16 50 250 50 50  32 44 30  90 25  62                                       71
35  12 19  36  4 30  18 48 26  68 36 130                                       71
30   9 22  48 31 15  28 46 45 202 34 116                                       74
60  36 18  32 13 46  22 52 29  94 27  73                                       74
DCNE
```

SAMPLE OUTPUT

```
12 INC. VAR.,GROUP1=SET1,GROUP2=SETS2-5,GROUP3=SETS6-7

CONSTANTS READ OR COMPUTED FROM CONTROL CARDS

         0=NTF (LABEL CARD FLAG,MUST BE ZERO OR ONE)
         1=NRF (SET CARD FLAG,MUST BE ZERO OR ONE)
         1=NGF (GROUP CARD FLAG,MUST BE ZERO OR ONE)
         0=NIF (INTERCEPT FLAG,MUST BE ZERO OR ONE)
         0=NFF (FORMAT CARD FLAG,MUST BE ZERO OR ONE)
        32=NOB (NUMBER OF OBSERVATIONS,MUST BE GREATER THAN THREE)
         7=NXS (NUMBER OF SETS OF INDEPENDENT VARIABLES NOT FIXED,MUST BE GREATER THAN ONE AND LESS THAN 31)
         1=NY  (NUMBER OF DEPENDENT VARIABLES,MUST BE LESS THAN NINE)
         0=NF  (NUMBER OF SETS OF FIXED INDEPENDENT VARIABLES,MUST BE LESS THAN 18)
         7=NSET(TOTAL NUMBER OF SETS OF INDEPENDENT VARIABLES,NF+NXS,MUST BE GREATER THAN ONE AND LESS THAN 31)
         3=LM  (LARGEST NUMBER OF SETS NOT FIXED TO BE INCLUDED IN ANY REGRESSION.MUST BE GREATER THAN ONE AND LESS THAN 18)
         3=NGX (NUMBER OF GROUPS,MUST BE EQUAL TO OR GREATER THAN LM AND LESS THAN OR EQUAL TO NXS)
        13=NVR (TOTAL NUMBER OF VARIABLES BEFORE TRANSFORMATIONS,MUST BE LESS THAN 51)
        13=NVS (TOTAL NUMBER OF VARIABLES AFTER TRANSFORMATIONS,MUST BE LESS THAN 51)
       29=NRGK (NUMBER OF REGRESSIONS,MUST BE LESS THAN 14,000 DIVIDED BY NY+1)

INPUT FORMAT
(3F3.1,F4.1,2F3.1,F4.1,2F3.1,F4.1,F3.1,F3.0,39X,F2.0)

DEPENDENT VARIABLE(S)
  Y1

SET LABELS AND NUMBER OF VARIABLES PER SET
   1   2   3   4   5   6   7
   4   1   1   2   1   1   2

NUMBER OF SETS PER GROUP,SUM MUST BE EQUAL TO NXS
   1   4   2

LISTING OF FIRST THREE SETS OF DATA

  0.1000000E 02  0.1000000E 01  0.2800000E 02  0.6800000E 02  0.4300000E 02  0.3400000E 02  0.2600000E 02  0.5000000E 02
  0.3100000E 02  0.9600000E 02  0.3500000E 02  0.1220000E 03  0.5100000E 02

  0.2000000E 02  0.4000000E 01  0.2000000E 02  0.4000000E 02  0.3100000E 02  0.1500000E 02  0.1800000E 02  0.5400000E 02
  0.1000000E 02  0.1000000E 02  0.5600000E 02  0.3140000E 03  0.5100000E 02

  0.1000000E 02  0.1000000E 01  0.2500000E 02  0.6200000E 02  0.5000000E 02  0.7500000E 02  0.3000000E 02  0.4600000E 02
  0.1600000E 02  0.2600000E 02  0.6600000E 02  0.4360000E 03  0.5300000E 02
LISTING OF LAST SET OF DATA
IF MISSING,PROGRAM HAS READ PAST DATA CARDS-CHECK NUMBER OF OBSERVATIONS,FORMAT CARD AND DATA CARDS

  0.6000000E 02  0.3600000E 02  0.1300000E 02  0.3200000E 02  0.1300000E 02  0.4600000E 02  0.2200000E 02  0.5200000E 02
  0.2900000E 02  0.8400000E 02  0.2700000E 02  0.7300000E 02  0.7400000E 02
```

continued

```
     CORRECTED SUM OF SQUARES

         Y1=  0.1923437E 04

12 IND. VAR.,GROUP1=SET1,GROUP2=SETS2-5,GROUP3=SETS6-7

REGRESSION SCREEN OUTPUT
COEFFICIENTS OF COLINEARITY AND DETERMINATION

                    IDENTIFICATION OF VARIABLES                       C OF C    Y1
 1                                                          1   0.43E-05 0.307
 1                                                          2   0.21E 00 0.083
 1                                                          3   0.19E 00 0.023
 1                                                          4   0.56E-03 0.014
 1                                                          5   0.17E 00 0.117
 1                                                          6   0.35E 00 0.088
 1                                                          7   0.14E-02 0.171

 2                                                      1   2   0.65E-06 0.367
 2                                                      1   3   0.69E-06 0.336
 2                                                      1   4   0.19E-08 0.336
 2                                                      1   5   0.64E-06 0.343
 2                                                      1   6   0.14E-05 0.335
 2                                                      1   7   0.55E-08 0.400
 2                                                      2   6   0.71E-01 0.139
 2                                                      2   7   0.29E-03 0.235
 2                                                      3   6   0.65E-01 0.096
 2                                                      3   7   0.20E-03 0.175
 2                                                      4   6   0.13E-03 0.120
 2                                                      4   7   0.76E-06 0.184
 2                                                      5   6   0.38E-02 0.137
 2                                                      5   7   0.14E-03 0.180

 3                                                  1   2   6   0.20E-06 0.379
 3                                                  1   2   7   0.82E-09 0.444
 3                                                  1   3   6   0.22E-06 0.354
 3                                                  1   3   7   0.65E-09 0.400
 3                                                  1   4   6   0.42E-09 0.433
 3                                                  1   4   7   0.24E-11 0.434
 3                                                  1   5   6   0.12E-07 0.349
 3                                                  1   5   7   0.46E-09 0.400


          ALL 12 INDEPENDENT VARIABLES                          0.17E-16 0.494

IHC9001 EXECUTION TERMINATING DUE TO ERROR COUNT FOR ERROR NUMBER  217

IHC2171 FIOCS - END OF DATA SET ON UNIT    1

TRACEBACK FOLLOWS-   ROUTINE    ISN   REG. 14    REG. 15    REG.  0    REG.  1

                     IBCOM            00060F60   00063B10   00000002   0004CF22

                     RDMNT     C004   4204C89C   000603A0   00000002   0004CF22

                     MAIN             0000C472   0104C568   FD000006   0006B7F8

ENTRY POINT=  C104C568

SUMMARY OF ERRORS FOR THIS JOB   ERROR NUMBER   NUMBER OF ERRORS

                                     217               1
```

COST ESTIMATE

For the job listed on the Sample Input/Output, the total running time was 1.50 seconds. At the current rate at Iowa State University, the chargeable computer time was $5.00 (Iowa State University minimum); postage and handling are included.

Charge to user = computer time + postage and handling + network overhead
= $5.00 (Iowa State minimum) + network overhead

CONTENTS—FURNIVAL

pages

1,2 Identification & Abstract
3–12 User Instructions
13–15 I/O
17 Cost—Contents

71

DESCRIPTIVE TITLE	FORTRAN Subroutine Package to Solve Ordinary Differential Equations
CALLING NAME	NODE
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Center, Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

NODE is a set of subroutines to solve systems of first-order ordinary differential equations, using discrete-variable methods. The predictor—corrector equations of Crane,[1] which have a wide range of stability, are implemented; the necessary backpoints are initially calculated by using the Runge—Katta—Gill single-step method.[2]

Systems for which NODE is adapted may arise in several ways. Theoretically, every ordinary differential equation of order higher than one can be reduced to a system of first-order equations. For a differential equation of order m, given by

$$y^{(m)} = f[x,y,y',y'',\cdots,y^{(m-1)}],$$

where f is a given function of its m+1 arguments. (Superscripts in parentheses denote derivatives.) The reduction to a first-order system is accomplished by setting

$$y_1 = y,\ y_2 = y',\ \cdots,\ y_m = y^{(m-1)}.$$

If the functions $y_1, y_2, \cdots, y_m$ satisfy the system

$$\begin{aligned} y_1' &= y_2 \\ y_2' &= y_3 \\ &\vdots \\ y_m' &= f(x,y_1,y_2,\cdots,y_m), \end{aligned}$$

continued

then the function $y(x) = y_1(x)$ evidently must satisfy the original equation.

Such first-order systems may also arise in a natural way from many physical problems.

REFERENCES

1. Crane, R.L., "Stability and Local Accuracy of Numerical Methods for Ordinary Differential Equations," Ph.D. thesis, Ia. State Univ. Sci. Technol. (unpublished).

2. Gill,S., "A Process for the Step-by-Step Integration of Differential Equations in an Automatic Digital Computing Machine," Proc. Cambridge Phil. Soc. *47*, 96 (1951).

"NODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969).

USER INSTRUCTIONS

To use the NODE package, a user must write a main program, which must be prefaced by the card

```
INTEGER ENDNO, ERRCK, CHANGH, T
```

and which should eventually call the NODE subroutine with

```
CALL NODE(VAR,H,N,X,ENDX,T,ENDNO,HRATIO,CHANGH,NODUM,ERRCK,IO)
```

The argument VAR is a real arraythat must be dimensioned in the user's main program as VAR(14,N), where N is the number of differential equations in the system, given explicitly in numerical form, e.g., DIMENSION VAR(14,5).

To use NODE, it is also necessary to initialize the following variables, either by reading from cards or by using FORTRAN statements in the user's main program.

Variable	*Initialization*
N	number of equations in the system
H	initial step size (see note below)
T	number of significant digits of accuracy needed (from 1 to 5)
NODUM	0: only "good" points are to be output 1: all points are to be output
ERRCK	0: absolute error check (see note below) 1: relative error check
HRATIO	s, where 0<s<1, if H, the step size, is to be reduced to sH (see note below)
CHANGH	0: suppress both reduction and doubling; i.e., do not change H under any condition 1: allow doubling but suppress reduction 2: allow reduction but suppress doubling 3: double step size when possible and reduce it according to HRATIO when necessary (see note below)
X	initial value of the independent variable x
ENDX	final value of the independent variable x
ENDNO	number of points in the interval of integration (if the endpoint ENDX is to be hit exactly). (In this case, H will be calculated automatically and HRATIO should be initialized to 0.5.

continued

Variable	*Initialization*
	If the endpoint procedure is not desired, ENDNO *must* be set to zero.)
VAR(1,I)	initial values of the dependent variables, where I ranges from 1 to N
IO	reference number of data set in which results are to be output

Note: The error at each point is approximated by the difference between the predicted value and the corrected value. The user has either of the following options:

When specifying an *absolute* error check—

$$E = \left| P_{n+1} - C_{n+1} \right|$$

When specifying a *relative* error check—

$$E = \left| \frac{P_{n+1} - C_{n+1}}{C_{n+1}} \right|$$

Included in the NODE routine is a user option to change the step size (H) of integration automatically, depending upon the size of the error estimate. The step H is set to sH (where s is specified by the user) if the largest value of E at a given step is greater than 16.219658 x 10**-T. Furthermore, the step size is doubled (2H) if for three (3) consecutive steps the maximum E is less than 0.08109829 x 10**-T. It is recommended that H be used as a power of 2 and that, if reduction is to be suppressed, HRATIO be set set to 1.0.

There are six FORTRAN subroutines called by NODE (not by the user's main program). The first three listed below, COMPD, COMPY, and COMPT, must be written by the user.

Subroutine	*Function*
COMPD	Computes the x-dependent parts of the derivatives y_i' and saves them for the COMPY subroutine. COMPD must then call COMPY
COMPY	Computes the derivatives y_i', using the previously computed x-dependent parts. Then these derivatives are to be stored in VAR(8,I)

continued

Subroutine	*Function*
COMPT	Outputs the results of the computation
PREDI	Predicts a new point with the predictor equation of Crane.[1] (The mathematical model of this equation is discussed in Ref. 2.)
CORRT	Corrects the point found by PREDI, using the corrector equation of Crane.[1] Then the computed values of the dependent variables are left in VAR(1,I)
INITA	Computes the initial backpoints required for this pair of predictor–corrector equations. (The mathematics of this procedure also is discussed in Ref 2.)

If there are no y-dependent parts to the derivatives, the subroutine COMPD will compute the entire derivative, and the subroutine COMPY will consist only of the cards

```
SUBROUTINE COMPY(VAR)
DIMENSION VAR(14,1)
RETURN
END
```

On the other hand, if the solution to a differential equation such as

$$y' = y \tan x + \sin x$$

is required, both COMPD and COMPY are used and may be coded as follows:

```
SUBROUTINE COMPD(VAR,X)
DIMENSION VAR(14,1)
COMMON STORE1, STORE2
STORE1 = SIN(X)
STORE2 = TAN(X)
CALL COMPY(VAR)
RETURN
END

SUBROUTINE COMPY(VAR)
DIMENSION VAR(14,1)
COMMON STORE1, STORE2
VAR(8,1) = VAR(1,1)*STORE2 + STORE1
RETURN
END
```

Note: Values may be conveniently carried from COMPD to COMPY by

continued

a variable or array in COMMON storage.

The output routine COMPT may be coded

```
      SUBROUTINE COMPT(VAR,N,X,IO)
      DIMENSION VAR(14,1)
      WRITE (IO,5) X, VAR(1,1), VAR(8,1)
    5 FORMAT (F10.3, 2E20.7)
      RETURN
      END
```

As another example, consider the third-order differential equation

$$x^3y''' + 6x^2y'' = 7xy' + y = 0$$

with the initial conditions $x = 1$, $y = -3$, $Y' = 1$, $y'' =1$. This differential equation reduces to the following system of ordinary differential equations.

$$y_1' = y_2$$

$$y_2' = y_3$$

$$y_3' = -(x^3)^{-1} \; [6x^2y_3 + 7xy_2 + y_1]$$

This system of differential equations is coded and solved in the Sample Input and Output.

REFERENCES

1. Crane, R.L., "Stability and Local Accuracy of Numerical Methods for Ordinary Differential Equations," Ph.D. thesis, Ia. State Univ. Sci Technol. (unpublished).

2. "NODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969).

SAMPLE INPUT/OUTPUT

```
FORTRAN IV G LEVEL 1, MOD 4             MAIN               DATE = 69195         18/49/54              PAGE 0001

 0001                DIMENSION VA3(14,5)
 0002                INTEGER ENDNO,T,ERRCK,CHANGH
             C       MAIN PROGRAM                                                          00020
             C
 00C3                X=1.
 0004                H=.025
 0005                AL=ALCG(X)
 0006                Y=(-3.-2.*AL+.5*AL*AL)/X
 0007             15 FORMAT(F10.5,2F15.4)
 0008                READ(1,1)H,N,X,ENDX,T,ENDNO,HRATIO,CHANGH,NODUM.ERRCK,IO
 0009              1 FORMAT(F5.0,I5,F5.0,F5.0,2I5,F5.0,4I5)
 0010                WRITE(3,3)H,N,X,ENDX,T,ENDNO,HRATIO,CHANGH,NODUM,ERRCK,IO
 0011              3 FORMAT('1USER SUPPLIED PARAMETERS',/,'    H= ',F9.3,/,
                    1 '    N= ',I9,/,'    X= ',F9.3,/,'    ENDX= ',F6.3,/,
                    2 '    T= ',I9,/,'    ENDNO= ',I5,/,'    HRATIO= ',F4.2,/,
                    3 '    CHANGH= ',I4,/,'    NODUM= ',I5,/,'    ERRCK= ',I5,/,
                    4 '    IU= ',I8)
 0012                VAR(1.1)=-3.
 0013                VAR(1,2)=1.
 0014                VAR(1,3)=1.
 0015                CALL  NODE(VAR,H,N,X,ENDX,T,ENDNO,HRATIO,CHANGH,NODUM,ERRCK,IO)
 0016                STOP                                                                  00190
 0017                END                                                                   00200
```

```
FORTRAN IV G LEVEL 1, MOD 4             MAIN               DATE = 69195         18/49/54              PAGE 0002

  TOTAL MEMORY REQUIREMENTS 000488 BYTES
```

continued

```
FORTRAN IV G LEVEL 1, MOD 4            COMPD              DATE = 69195          18/49/54               PAGE 0001

 0001                 SUBROUTINE  COMPD(VAR,X)
            C     USER'S RESPONSIBILITY
            C  CALCULATE DERIVATIVE - (X-DEPENDENT ALWAYS )
            C
 0002                 DIMENSION VAR(14,1),STORE(5)
 0003                 CCMMON STORE
 0004                 STORE(1)=-1.00/( X*X*X)
 0005                 STORE(2)=6.00*X*X
 0006                 STORE(3)=7.00*X
 0007                 CALL   COMPY(VAR)
 0008                 RETURN                                                              00320
 0009                 END                                                                 00330


FORTRAN IV G LEVEL 1, MOD 4            COMPD              DATE = 69195          18/49/54               PAGE 0002

  TOTAL MEMORY REQUIREMENTS 00018E BYTES




FORTRAN IV G LEVEL 1, MOD 4            CCMPY              DATE = 69195          18/49/54               PAGE 0001

 0001                 SUBROUTINE  COMPY(VAR)
            C     USER'S RESPONSIBILITY
            C  CALCULATE DERIVATIVE WHEN COMPD CALCULATES X-DEPENDENT PARTS
            C
 0002                 DIMENSION VAR(14,1),STORE(5)
 0003                 COMMON STORE
 0004                 VAR(8,1)=VAR(1,2)
 0005                 VAR(8,2)=VAR(1,3)
 0006                 VAR(8,3)=STORE(1) * (STORE(2)*VAR(1,3)+STORE(3)*VAR(1,2)+VAR(1,1))
 0007                 RETURN                                                              00370
 0008                 END                                                                 00380


FORTRAN IV G LEVEL 1, MOD 4            COMPY              DATE = 69195          18/49/54               PAGE 0002

  TOTAL MEMORY REQUIREMENTS 000140 BYTES


FORTRAN IV G LEVEL 1, MOD 4            CCMPT              DATE = 69195          18/49/54               PAGE 0001

 0001                 SUBROUTINE  COMPT(VAR,N,X,IO)
            C
            C  WRITE-OUT   ROUTINE  -- USER'S RESPONSIBILITY
            C
 0002                 DIMENSION VAR(14,1)
 0003               8 WRITE (IO,1) X, (VAR(1,I)          ,I = 1,N )
 0004               1 FORMAT ( F8.4 , 3F15.4)
 0005                 RETURN
 0006                 END                                                                 00680


FORTRAN IV G LEVEL 1, MOD 4            COMPT              DATE = 69195          18/49/54               PAGE 0002

  TOTAL MEMORY REQUIREMENTS 0001CC BYTES
```

continued

```
USER SUPPLIED PARAMETERS
  H=      0.025
  N=          3
  X=      1.000
  ENDX=   1.500
  T=          4
  ENDNO=     80
  HRATIO= 0.20
  CHANGH=     3
  NODUM=      0
  ERRCK=      1
  IO=         3




 1.0000        -3.0000         1.0000         1.0000
 1.0062        -2.9937         1.0061         0.9387
 1.0125        -2.9874         1.0117         0.8798
 1.0187        -2.9811         1.0171         0.8231
 1.0250        -2.9747         1.0220         0.7687
 1.0312        -2.9683         1.0267         0.7163
 1.0375        -2.9619         1.0310         0.6659

          TEPSIZE DOUBLED
 1.0500        -2.9489         1.0387         0.5739
 1.0625        -2.9359         1.0453         0.4830
 1.0750        -2.9228         1.0508         0.4017

         STEPSIZE DOUBLED
 1.1000        -2.8964         1.0590         0.2569
 1.1250        -2.8699         1.0638         0.1330

         STEPSIZE REDUCED
 1.1375        -2.8566         1.0651         0.0780
 1.1500        -2.8433         1.0658         0.0271
 1.1625        -2.8299         1.0658        -0.0230
 1.1750        -2.8166         1.0653        -0.0634
 1.1875        -2.8033         1.0643        -0.1036
 1.2000        -2.7900         1.0627        -0.1405
 1.2125        -2.7767         1.0607        -0.1748
 1.2250        -2.7635         1.0584        -0.2064

         STEPSIZE DOUBLED
 1.2500        -2.7371         1.0525        -0.2623
 1.2750        -2.7109         1.0453        -0.3096
 1.3000        -2.6848         1.0371        -0.3494
 1.3250        -2.6590         1.0279        -0.3829
 1.3500        -2.6335         1.0180        -0.4108
 1.3750        -2.6081         1.0074        -0.4338
 1.4000        -2.5831         C.9963        -0.4527
 1.4250        -2.5583         0.9848        -0.4678
 1.4500        -2.5338         0.9729        -0.4799
 1.4750        -2.5097         0.9608        -0.4891
 1.5000        -2.4858         0.9485        -0.4960
 1.5250        -2.4622         0.9360        -0.5007
```

COST ESTIMATE

For the job listed on the Sample Input/Output, the total running time was 0.30 seconds. At the current rate at Iowa State University, the chargeable computer time was $5.00 (Iowa State University minimum); postage and handling are included.

Charge to user = computer time + postage and handling + network overhead

= $5.00 (Iowa State minimum) + network overhead

CONTENTS—NODE

pages	
1,2	Identification & Abstract
3–6	User Instructions
7–9	I/O
11	Cost—Contents

DESCRIPTIVE TITLE	FORTRAN Subroutine Package to Solve Ordinary Differential Equations
CALLING NAME	DNODE
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Claire G. Maple, Director, Computation Center, Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

DNODE is a set of subroutines to solve systems of first-order ordinary differential equations, using discrete-variable methods and double-precision arithmetic. The method of computation is identical to that of the NODE package (EIN No. 000 0082); however, the implementation of double-precision necessitates a slight revision in parameter initialization of the NODE subroutines. (Please refer to the NODE package description.)

REFERENCES

"DNODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969).

"NODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969). See also the EIN entry for NODE (EIN No. 000 0082).

Crane, R.L., "Stability and Local Accuracy of Numerical Methods for Ordinary Differential Equations," Ph.D. thesis, Ia. State Univ. Sci. Technol. (unpublished).

USER INSTRUCTIONS

Usage of the DNODE package is essentially the same as for the NODE package (EIN No. 000 0082), except for the following exceptions.

1. The user's main program must now be prefaced by the cards

```
INTEGER ENDNO, ERRCK, CHANGH, T
DOUBLE PRECISION VAR(14,N), X, ENDX, H, HRATIO
```

   where N is the number of differential equations in the system, given explicitly in numerical form, e.g.,

```
DOUBLE PRECISION VAR(14,5), X, ENDX, H, HRATIO
```

2. In initializing HRATIO, the value should be set to 1.0D0, *if reduction is to be suppressed.* (All other parameters are initialized as in the NODE package.)

3. The corresponding six subroutines in the DNODE package are DCOMPD, DCOMPY, DCOMPT, DPREDI, DCORRT, and DINITA. The purpose of each subroutine is identical to that of its counterpart in the NODE package, and, as in NODE, the first three (DCOMPD, DCOMPY, and DCOMPT) must be written by the user. The subroutine DCOMPD must call the subroutine DCOMPY.

4. If there are no y-dependent parts to the derivatives, the subroutine DCOMPD will compute the entire derivative and the subroutine DCOMPY will consist only of the cards

```
SUBROUTINE DCOMPY
DOUBLE PRECISION VAR(14,1)
RETURN
END
```

5. For the given example in the NODE package—namely, the solution of the differential equation

$$y' = y \tan x + \sin x$$

   —both DCOMPD and DCOMPY are used and may be coded as follows.

```
SUBROUTINE DCOMPD(VAR,X)
DOUBLE PRECISION VAR(14,1), X, STORE1, STORE2
COMMON STORE1, STORE 2
STORE1 = DSIN(X)
STORE2 = DTAN(X)
CALL DCOMPY(VAR)
RETURN
END
```

continued

```
      SUBROUTINE DCOMPY(VAR)
      DOUBLE PRECISION VAR(14,1), X, STORE1, STORE2
      COMMON STORE1, STORE2
      VAR(8,1) = VAR(1,1)*STORE2 + STORE1
      RETURN
      END
```

Note: Values may be conveniently carried from DCOMPD to DCOMPY by a variable or array in COMMON storage.

6. The output routine may be coded

```
      SUBROUTINE DCOMPT(VAR,N,X,IO)
      DOUBLE PRECISION VAR(14,1), X
      WRITE (IO,5) X, VAR(1,1), VAR(8,1)
    5 FORMAT (F10.2, 2D25.14)
      RETURN
      END
```

7. In the user's main program and subroutines DCOMPD, DCOMPY, and DCOMPT, any variable or constant that ~~is~~ used in computation *must* be declared DOUBLE PRECISION.

For purposes of illustration, the same example as used in the NODE description—namely, the differential equation

$$x^3y''' + 6x^2y'' + 7xy' + y = 0,$$

with initial conditions $x = 1$, $y = -3$, $y' = 1$, $y'' = 1$—is coded and solved in the Sample Input and Output (in its equivalent form as a system of first-order equations).

REFERENCES

"DNODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969).

"NODE: A Package of Fortran Subroutines to Solve Ordinary Differential Equations" (user manual), Ia. State Univ. Computation Ctr. (July 1969).

SAMPLE INPUT/OUTPUT

```
FORTRAN IV G LEVEL 1, MOD 4            MAIN              DATE = 69191         11/37/37            PAGE 0001

 0001                DOUBLE PRECISION VAR(14,5),H,X,ENDX,HRATIO,AL,Y
 0002                INTEGER ENDNO,T,ERRCK,CHANGH
             C       MAIN PROGRAM                                                         00020
             C
 0003                X=1.D0
 0004                H=.025D0
 0005                AL=DLOG(X)
 0006                Y=(-3.D0 - 2.D0*AL + .5D0*AL*AL)/X
 0007             15 FORMAT(F10.5,2D25.14)
 0008                READ(1,1)H,N,X,ENDX,T,ENDNO,HRATIO,CHANGH,NODUM,ERRCK,IO
 0009              1 FORMAT(F5.0,I5,F5.0,F5.0,2I5,F5.0,4I5)
 0010                WRITE(3,3)H,N,X,ENDX,T,ENDNO,HRATIO,CHANGH,NODUM,ERRCK,IO
 0011              3 FORMAT('1USER SUPPLIED PARAMETERS',/,'   H= ',F9.3,/,
                    1 '   N= ',I9,/,'   X= ',F9.3,/,'   ENDX= ',F6.3,/,
                    2 '   T= ',I9,/,'   ENDNO= ',I5,/,'   HRATIO= ',F4.2,/,
                    3 '   CHANGH= ',I4,/,'   NODUM= ',I5,/,'   ERRCK= ',I5,/,
                    4 '   IO= ',I8)
 0012                VAR(1,1)=-3.D0
 0013                VAR(1,2)=1.D0
 0014                VAR(1,3)=1.D0
 0015                TIME=0.
 0016                CALL DNODE(VAR,H,N,X,ENDX,T,ENDNO,HRATIO,CHANGH,NODUM,ERRCK,IO)
 0017                STOP                                                                 00190
 0018                END                                                                  00200
```

```
FORTRAN IV G LEVEL 1, MOD 4            MAIN              DATE = 69191         11/37/37            PAGE 0002

  TOTAL MEMORY REQUIREMENTS 000608 BYTES
```

continued

```
FORTRAN IV G LEVEL 1, MOD 4           DCOMPD             DATE = 69191         11/37/37                PAGE 0001

 0001                SUBROUTINE DCCMPD(VAR,X)
           C         USER'S RESPONSIBILITY
           C   CALCULATE DERIVATIVE - (X-DEPENDENT ALWAYS )
           C
 0002                   DOUBLE PRECISION VAR(14,1),X,ENDX,STORE(5)
 0003                COMMON STORE
 0004                STORE(1)=-1.D0/(X*X*X)
 0005                STORE(2)=6.D0*X*X
 0006                STORE(3)=7.D0*X
 0007                CALL DCOMPY(VAR)
 0008                RETURN                                                          00320
 0009                END                                                             00330


FORTRAN IV G LEVEL 1, MOD 4           DCOMPD             DATE = 69191         11/37/37                PAGE 0002

  TOTAL MEMORY REQUIREMENTS 0001A2 BYTES


FORTRAN IV G LEVEL 1, MOD 4           DCOMPY             DATE = 69191         11/37/37                PAGE 0001

 0001                SUBROUTINE DCCMPY(VAR)
           C         USER'S RESPONSIBILITY
           C   CALCULATE DERIVATIVE WHEN COMPD CALCULATES X-DEPENDENT PARTS
           C
 0002                   DOUBLE PRECISION VAR(14,1),X,ENDX,STORE(5)
 0003                COMMON STORE
 0004                VAR(8,1)=VAR(1,2)
 0005                VAR(8,2)=VAR(1,3)
 0006                VAR(8,3)=STCRE(1) * (STORE(2)*VAR(1,3)+STORE(3)*VAR(1,2)+VAR(1,1))
 0007                RETURN                                                          00370
 0008                END                                                             00380


FORTRAN IV G LEVEL 1, MCD 4           DCOMPY             DATE = 69191         11/37/37                PAGE 0002

  TOTAL MEMORY REQUIREMENTS 000140 BYTES


FORTRAN IV G LEVEL 1, MOD 4           DCOMPT             DATE = 69191         11/37/37                PAGE 0001

 0001                SUBROUTINE DCOMPT(VAR,N,X,IO)
           C
           C   WRITE-OUT   ROUTINE  -- USER'S RESPONSIBILITY
           C
 0002                   DOUBLE PRECISION VAR(14,1),X,ENDX
 0003             8 WRITE (IO,1) X, (VAR(1,I)         ,I = 1,N )
 0004             1 FORMAT ( F8.4 , 3D24.14)
 0005                RETURN
 0006                END                                                             00680


FORTRAN IV G LEVEL 1, MOD 4           DCOMPT             DATE = 69191         11/37/37                PAGE 0002

  TOTAL MEMORY REQUIREMENTS 000104 BYTES
```

continued

```
USER SUPPLIED PARAMETERS
  H=       0.025
  N=           3
  X=       1.000
  ENDX=    1.500
  T=          10
  ENDNO=      80
  HRATIO=  0.20
  CHANGH=      3
  NODUM=       0
  ERRCK=       1
  IO=          3


  1.0000   -0.30000000000000D 01    0.10000000000000D 01    0.10000000000000D 01

           STEPSIZE REDUCED
  1.0031   -0.29968701678001D 01    0.10030764906723D 01    0.96905552840429D 00
  1.0062   -0.29937308716768D 01    0.10060572249476D 01    0.93871385208920D 00
  1.0094   -0.29905824079454D 01    0.10089440674511D 01    0.90896278645712D 00
  1.0125   -0.29874250671576D 01    0.10117388451855D 01    0.87979041807752D 00
  1.0156   -0.29842591342048D 01    0.10144433482812D 01    0.85118509859980D 00
    .            .                        .                       .
    .            .                        .                       .
    .            .                        .                       .
  1.1250   -0.28698930218963D 01    0.10638322454811D 01    0.13301576632236D 00
  1.1281   -0.28665679198097D 01    0.10642257126733D 01    0.11884941637277D 00
  1.1312   -0.28632416568608D 01    0.10645753402938D 01    0.10495757226143D 00

           STEPSIZE REDUCED
  1.1328   -0.28615781325614D 01    0.10647339804321D 01    0.98112974247527D-01
  1.1344   -0.28599143687149D 01    0.10648819780782D 01    0.91335077048662D-01
  1.1359   -0.28582503818688D 01    0.10650194369604D 01    0.84623252377561D-01
  1.1375   -0.28565861884098D 01    0.10651464598308D 01    0.77976878341066D-01
  1.1391   -0.28549218045644D 01    0.10652631484740D 01    0.71395339374678D-01
    .            .                        .                       .
    .            .                        .                       .
    .            .                        .                       .
  1.2437   -0.27436956384936D 01    0.10541022409910D 01   -0.24913841980310D 00
  1.2453   -0.27420489123759D 01    0.10537083395420D 01   -0.25246807089465D 00
  1.2469   -0.27404028026283D 01    0.10533112792428D 01   -0.25576342126070D 00



           STEPSIZE DOUBLED
  1.2500   -0.27371124643426D 01    0.10525018443563D 01   -0.26225245115988D 00
  1.2531   -0.27338246870073D 01    0.10516723405322D 01   -0.26860794299814D 00
  1.2562   -0.27305395326897D 01    0.10508231813346D 01   -0.27483228439749D 00
  1.2594   -0.27272570621763D 01    0.10499547729364D 01   -0.28092781802090D 00
  1.2625   -0.27239773349957D 01    0.10490675142590D 01   -0.28689684249659D 00
    .            .                        .                       .
    .            .                        .                       .
    .            .                        .                       .
  1.4937   -0.24917574876124D 01    0.95162590944269D 00   -0.49444963192671D 00
  1.4969   -0.24887860722173D 01    0.95007954247142D 00   -0.49521967553354D 00
  1.5000   -0.24858194929250D 01    0.94853082088421D 00   -0.49595668352480D 00
  1.5031   -0.24828577569332D 01    0.94697984700933D 00   -0.49666123539914D 00
```

COST ESTIMATE

For the job listed on the Sample Input/Output, the total running time was 1.20 seconds. At the current rate at Iowa State University, the chargeable computer time was $5.00 (Iowa State University minimum); postage and handling are included.

Charge to user = computer time + postage and handling + network overhead

= $5.00 (Iowa State minimum) + network overhead

CONTENTS—DNODE

pages

1 Identification & Abstract
3,4 User Instructions
5–7 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Bucharest Sort a List into Ascending Order
CALLING NAME	ABSRT
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

ABSRT is a subroutine to sort the contents of an array into ascending order [defined as $A(N) \leq A(N+1)$]. It is written for a list of real numbers; however, with a slight revision in a declaration statement within the program deck, it may be easily altered to sort a list of integers or double-precision real numbers.

A formal description of the two-way, internal, single-phase merge method may be found in the literature.[1] Briefly described, the procedure considers the input list to be composed of a number of shorter sublists that are already ordered. (Initially, every element in the list comprises a sublist.) Half of the sublists are stored in standard order, forward from A(1), and the other half are stored backwards from A(N)—i.e., in A(N), A(N–1), A(N–2), etc. The first pass through the input list merges pairs of these sublists, taking one from each end of the input list, and stores each newly merged pair in the output list. The output list is stored in A(N+1) through A(2N), and, like the input list, is made up of sublists stored forward from A(N+1) and backward from A(2N). On successive passes through the list, the input and output areas are switched, and consecutively, the number of sublists is halved. When only one sublist remains, the list is in order. At that time, a check is made to determine if the sorted list is in the first or second half of the array; if necessary, a final pass is made to put the list back into the first half, A(1) through A(N).

REFERENCES

1. Iverson, K.E., *A Programming Language* (John Wiley & Sons, Inc., New York, 1962), pp.206–211.

USER INSTRUCTIONS

The calling sequence for this subroutine is

```
CALL ABSRT(N,A)
```

with the two arguments supplied as follows.

N	an integer equal to the number of elements to be sorted (not altered by the routine)
A	a singly subscripted array containing both the list of elements to be sorted and a temporary work area. ABSRT *must* be entered with the list to be sorted in the first N locations of the array A. The first 2N+1 locations of A will be destroyed by the subroutine, and the final sorted list will replace the original unsorted one. Note that, to avoid error, the array A must be dimensioned at least 2N+1.

Of course, to make use of the subroutine ABSRT, a main program, which will, in turn, make a call to ABSRT, must be supplied by the user. Such details as format (for reading, printing, punching, etc.) should be contained therein (as well as the dimension statement for the array to be sorted). Such a program is listed in the Sample Input.

Note: As mentioned in the Abstract, the alteration of ABSRT to a subroutine for the sorting of integer or double-precision lists is quite simple. The only card that requires a change is a declaration statement (card ABSRT035); however, it should be realized that sorts on other modes will require different amounts of time, with integer sorts taking less time and double-precision sorts taking more time, on the average.

REFERENCES

"ABSRT: Bucharest Sort a List into Ascending Order" (user manual), Ia. State Univ. Sci. Technol. (July 1969).

SAMPLE INPUT

```
C
C     SAMPLE CALLING PROGRAM FOR SUBROUTINE ABSRT
C
      DIMENSION ARRAY(21)
      READ(1,100)(ARRAY(I),I=1,10)
  100 FORMAT(10F5.1)
      WRITE(3,200)(ARRAY(I),I=1,10)
  200 FORMAT(15H1ORIGINAL LIST=,10F6.1)
C
      CALL ABSRT(10,ARRAY)
C
      WRITE(3,300)(ARRAY(I),I=1,10)
  300 FORMAT(15H0  SORTED LIST=,10F6.1)
      STOP
      END
```

SAMPLE OUTPUT

```
 FORTRAN IV G LEVEL 1, MOD 4            MAIN              DATE = 69191        11/34/53              PAGE 0001

              C
              C      SAMPLE CALLING PROGRAM FOR SUBROUTINE ABSRT
              C
 0001                DIMENSION ARRAY(21)
 0002                READ(1,100)(ARRAY(I),I=1,10)
 0003            100 FORMAT(10F5.1)
 0004                WRITE(3,200)(ARRAY(I),I=1,10)
 0005            200 FORMAT(15H1ORIGINAL LIST=,10F6.1)
              C
 0006                CALL ABSRT(10,ARRAY)
              C
 0007                WRITE(3,300)(ARRAY(I),I=1,10)
 0008            300 FORMAT(15H0  SORTED LIST=,10F6.1)
 0009                STOP
 0010                END


 FORTRAN IV G LEVEL 1, MOD 4            MAIN              DATE = 69191        11/34/53              PAGE 0002

   TOTAL MEMORY REQUIREMENTS 00027E BYTES




ORIGINAL LIST=  -1.0   5.1  -0.7   1.2   3.5   0.0  -0.5   0.5  -1.6   2.0

  SORTED LIST=  -1.6  -1.0  -0.7  -0.5   0.0   0.5   1.2   2.0   3.5   5.1
```

FILMED FROM BEST AVAILABLE COPY

COST ESTIMATE

The execution time for sorting a list is proportional to CN $\log_2$ N, where N is the length of the list to be sorted and C is a constant that depends on the machine and compiler used, the original order of the list, and the mode of the list. A worst case of 1000 real numbers was sorted under E-level FORTRAN on the IBM 360/50 in less than 2 seconds.

However, for the job listed on the Sample Input, the total running time was 0.20 seconds. At the current rate at Iowa State University, the chargeable computer time was $5.00 (Iowa State University minimum); postage and handling are included.

Charge to user = computer time + postage and handling + network overhead

= $5.00 (Iowa State minimum) + network overhead

CONTENTS—ABSRT

pages

1 Identification & Abstract
3 User Instructions
5,6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Bucharest Sort a List into Descending Order
CALLING NAME	DBSRT
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

DBSRT is a subroutine to sort the contents of an array into descending order [defined as $A(N) \geq A(N+1)$]. It is written for a list of real numbers; however, with a slight revision in a declaration statement within the program deck, it may be easily altered to sort a list of integers or double-precision real numbers.

The method used is identical to that described in the entry for subroutine ABSRT (EIN No. 000 0084), except, of course, for the reversed ordering of the elements within the mentioned sublists. For convenience, the reference from which a formal description of the sorting algorithm (the two-way, internal, single-phase merge method) may be obtained is listed below.

REFERENCES

Iverson, K.E., *A Programming Language* (John Wiley & Sons, Inc., New York, 1962), pp. 206–211.

USER INSTRUCTIONS

The calling sequence for this subroutine is

```
CALL DBSRT(N,A)
```

with the two arguments supplied as follows.

N	an integer equal to the number of elements to be sorted (not altered by the routine)
A	a singly subscripted array containing both the list of elements to be sorted and a temporary work area. DBSRT *must* be entered with the list to be sorted in the first N locations of the array A. The first 2N+1 locations of A will be destroyed by the subroutine, and the final sorted list will replace the original unsorted one. Note that, to avoid error, the array A must be dimensioned at least 2N+1.

Of course, to make use of the subroutine DBSRT (or ABSRT, EIN No. 000 0084), the user must supply a main program, which will, in turn, make a call to DBSRT. Such details as format (for reading, printing, punching, etc.) should be contained therein (as well as the dimension statement for the array to be sorted). Such a program is listed in the Sample Input.

Note: As mentioned in the Abstract, the alteration of DBSRT to a subroutine for the sorting of integer or double-precision lists is quite simple. The only card in the program deck that requires a change is a declaration statement (card DBSRT035); however, it should be realized that sorts on other modes will require different amounts of time, with integer sorts taking less time and double-precision sorts taking more time, on the average.

REFERENCES

"DBSRT: Bucharest Sort a List into Descending Order" (user manual), Ia. State Univ. Sci. Technol. (July 1969).

"ABSRT: Bucharest Sort a List into Ascending Order" (user manual), Ia. State Univ. Sci. Technol. (July 1969). See also the entry for ABSRT (EIN No. 000 0084).

SAMPLE INPUT

```
C
C     SAMPLE CALLING PROGRAM FOR SUBROUTINE DBSRT
C
      DIMENSION ARRAY(21)
      READ(1,100)(ARRAY(I),I=1,10)
  100 FORMAT(10F5.1)
      WRITE(3,200)(ARRAY(I),I=1,10)
  200 FORMAT(15H1ORIGINAL LIST=,10F6.1)
C
      CALL DBSRT(10,ARRAY)
C
      WRITE(3,300)(ARRAY(I),I=1,10)
  300 FORMAT(15H0  SORTED LIST=,10F6.1)
      STOP
      END
```

FILMED FROM BEST AVAILABLE COPY

SAMPLE OUTPUT

```
 FORTRAN IV G LEVEL 1, MOD 4            MAIN               DATE = 69191         11/33/53              PAGE 0001

              C
              C      SAMPLE CALLING PROGRAM FOR SUBROUTINE DBSRT
              C
 0001                DIMENSION ARRAY(21)
 0002                READ(1,100)(ARRAY(I),I=1,10)
 0003            100 FORMAT(10F5.1)
 0004                WRITE(3,200)(ARRAY(I),I=1,10)
 0005            200 FORMAT(15H0ORIGINAL LIST=,10F6.1)
              C
 0006                CALL DBSRT(10,ARRAY)
              C
 0007                WRITE(3,300)(ARRAY(I),I=1,10)
 0008            300 FORMAT(15H0  SORTED LIST=,10F6.1)
 0009                STOP
 0010                END


 FORTRAN IV G LEVEL 1, MOD 4            MAIN               DATE = 69191         11/33/53              PAGE 0002

   TOTAL MEMORY REQUIREMENTS 00027E BYTES




ORIGINAL LIST=  -1.0   5.1  -0.7   1.2   3.5   0.0  -0.5   0.5  -1.6   2.0

  SORTED LIST=   5.1   3.5   2.0   1.2   0.5   0.0  -0.5  -0.7  -1.0  -1.6
```

COST ESTIMATE

The execution time for sorting a list is proportional to CN $\log_2 N$, where N is the length of the list to be sorted and C is a constant that depends on the machine and compiler used, the original order of the list, and the mode of the list. A worst case of 1000 real numbers was sorted under E-level FORTRAN on the IBM 360/50 in less than 2 seconds.

However, for the job listed on the Sample Input, the total running time was 0.20 seconds. At the current rate at Iowa State University, the chargeable computer time was $5.00 (Iowa State University minimum); postage and handling are included.

Charge to user = computer time + postage and handling + network overhead
= $5.00 (Iowa State minimum) + network overhead

CONTENTS—DBSRT

pages

1 Identification & Abstract
3 User Instructions
5,6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	FORTRAN Subroutine to Solve Simultaneous Linear Equations with Complex Coefficients
CALLING NAME	CGELG
INSTALLATION NAME	Iowa State University Computation Center
AUTHOR(S) AND AFFILIATION(S)	Iowa State University Computation Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM System 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Dr. Clair G. Maple, Director, Computation Ctr., Iowa State University, Ames, Ia. 50010 Tel.: (515) 294-3402

FUNCTIONAL ABSTRACT

CGELG is a FORTRAN subroutine that solves a set of simultaneous equations with complex coefficients. There exists the capability of handling more than one set of righthand constraints at a time. Yet, the most desirable aspect of the routine is its high degree of accuracy, even on a large, ill-conditioned matrix.

The procedure used is the Gaussian elimination (with complete pivoting) method,[1] modified for complex arithmetic. In this method, the first step is a forward reduction, in which the coefficient matrix is reduced to an upper triangular matrix, using elementary row and column transformations. Then the solution is found by successive back substitution.

REFERENCES

1. Ralston, A., *A First Course in Numerical Analysis* (McGraw-Hill Book Co., Inc., New York, 1965), pp. 339–407.

USER INSTRUCTIONS

Description of Input

The calling sequence for CGELG is

```
CALL CGELG(R,A,M,MM,N,EPS,IER)
```

where the seven arguments are supplied as follows.

R	M by N complex matrix of righthand sides (original contents destroyed and replaced by solution)
A	M by M complex matrix of coefficients, stored by columns (original contents destroyed)
M	number of equations in the system to be solved
MM	dimension of first parameter of the A and R matrices (maximum number of equations)
N	number of sets of righthand constraints
EPS	real constant to be used as relative tolerance for a test on loss significance (see Error Information below)

Of course, to make use of CGELG, the user must supply a main program that will, in turn, make a call to CGELG. Such a program should allow for the reading of input data (matrices and parameters) or set the same; it should also arrange for the manipulation of output (printing, punching, etc.) as returned by CGELG. Such a program, used to call CGELG to solve a set of six equations, with four sets of righthand constraints, is presented in the Sample Input.

Error Information

At each stage in the matrix reduction, CGELG selects the absolutely largest eligible element as the pivot element. If the magnitude of the pivot element so chosen is exactly zero, then the set of equations is singular and IER will be set to -1. If the magnitude of the pivot element is very small, the equations are either singular or poorly conditioned. If a small pivot (less than or equal to EPS, times the largest element in matrix A) is found at reduction step K+1, then IER is set equal to K. (If A is well scaled and EPS is appropriately chosen, IER=K may be taken to indicate that matrix A is of rank K.)

continued

In summary,

Error Code	*Meaning*
0	no error
–1	no result ($M < 1$ or determinant = 0)
K	loss of significance at elimination step K+1, because pivot element ≤ EPS X largest element of A

REFERENCES

"CGELG: Fortran Subroutine to Solve Simultaneous Linear Equations with Complex Coefficients" (user manual), Ia. State Univ. Sci. Technol. (July 1969).

SAMPLE INPUT

```
C
C     SAMPLE MAIN PROGRAM TO CALL SUBROUTINE CGELG
C
      COMPLEX AA(10,10), A(100), RR(10,4), R(10)
      EQUIVALENCE (A,AA), (R,RR)
      M = 6
      N = 4
      READ(1,1) ((AA(I,J), J = 1,M), I = 1,M)
    1 FORMAT(12F5.0)
      WRITE(3,4) ((AA(I,J), J = 1,M), I = 1,M)
    4 FORMAT('1COEFFICIENT MATRIX',/,(' ', 6(2F5.1,10X)))
      READ(1,2) ((RR(I,J), J = 1,N), I = 1,M)
    2 FORMAT(8F5.0)
      WRITE(3,5) ((RR(I,J), J = 1,N), I = 1,M)
    5 FORMAT('0RIGHT-HAND SIDE MATRIX',/,(' ', 6(2F5.1,10X)))
C
      CALL CGELG(R,A,M,10,N,.0001,IER)
C
      WRITE(3,3) IER, ((RR(I,J), J = 1,N), I = 1,M)
    3 FORMAT('0IER=', I5,/,'0SOLUTION MATRIX',/,(' ', 4(2F7.1,10X)))
      STOP
      END
```

SAMPLE OUTPUT

```
FORTRAN IV G LEVEL 1, MOD 4             MAIN             DATE = 69191        11/35/51             PAGE 0001

               C
               C     SAMPLE MAIN PROGRAM TO CALL SUBROUTINE CGELG
               C
 0001                COMPLEX AA(10,10),A(100),RR(10,4),R(10)
 0002                EQUIVALENCE (A,AA),(R,RR)
 0003                M=6
 0004                N=4
 0005                READ(1,1) ((AA(I,J),J=1,M),I=1,M)
 0006              1 FORMAT(12F5.0)
 0007                WRITE(3,4)((AA(I,J),J=1,M),I=1,M)
 0008              4 FORMAT('1COEFFICIENT MATRIX',/,(' ',6(2F5.1,10X)))
 0009                READ(1,2) ((RR(I,J),J=1,N),I=1,M)
 0010              2 FORMAT(8F5.0)
 0011                WRITE(3,5)((RR(I,J),J=1,N),I=1,M)
 0012              5 FORMAT('0RIGHT-HAND SIDE MATRIX',/,(' ',6(2F5.1,10X)))
               C
 0013                CALL CGELG(R,A,M,10,N,.0001,IER)
               C
 0014                WRITE(3,3) IER,((RR(I,J),J=1,N),I=1,M)
 0015              3 FORMAT('0IER=',I5,/,'0SOLUTION MATRIX',/,(' ',4(2F7.1,10X)))
 0016                STOP
 0017                END

FORTRAN IV G LEVEL 1, MOD 4             MAIN             DATE = 69191        11/35/51             PAGE 0002

  TOTAL MEMORY REQUIREMENTS 00C892 BYTES



COEFFICIENT MATRIX
  1.0  1.0            1.0  1.0            2.0  0.0            4.0  0.0            0.0  1.0            0.0 -1.0
  0.0  2.0            2.0  2.0            1.0  0.0            2.0  0.0            0.0 -1.0            1.0  0.0
  1.0  0.0            1.0  0.0            1.0  0.0            1.0  0.0            1.0  0.0            1.0  0.0
  2.0  0.0            3.0  0.0            0.0  0.0            5.0  0.0            0.0  0.0            1.0  0.0
  0.0  2.0            0.0  2.0            0.0  3.0            0.0  6.0            0.0  5.0            0.0  4.0
  1.0  1.0            1.0  2.0            1.0  3.0            1.0  4.0            1.0  5.0            1.0  6.0

RIGHT-HAND SIDE MATRIX
  5.0  3.0            8.0  0.0            8.0  2.0           -1.0  0.0            3.0  3.0            2.0 14.0
  6.0  3.0            1.0  2.0            4.0  2.0            8.0  3.0            6.0  0.0            0.0  1.0
  8.0  4.0           15.0  9.0           11.0  0.0            0.0  1.0           -5.0 14.0           -6.0 24.0
  0.0 22.0           -5.0  0.0            2.0 14.0            2.0 30.0            6.0 21.0           -6.0  1.0

IER=    0

SOLUTION MATRIX
    0.0    1.0              3.0   -0.0              1.0   -0.0             -0.0   -1.0
    1.0    1.0             -0.0    3.0              1.0    0.0              0.0    1.0
    2.0   -0.0              2.0    0.0              1.0    0.0             -0.0    0.0
    1.0    0.0              1.0   -0.0              1.0   -0.0             -0.0   -0.0
    0.0    1.0             -2.0    0.0              1.0    0.0              0.0    1.0
   -0.0   -1.0              4.0   -0.0              1.0   -0.0             -0.0    0.0
```

COST ESTIMATE

The execution time for solving a set of M equations is proportional to CM^3, where C is a constant that depends on the machine and compiler used and the mode and conditioning of the coefficient matrix.

For the job listed, compiled and run under G-level FORTRAN on the IBM 360/65, the total time amounted to 0.40 seconds. At the current rate at Iowa State University, the chargeable computer time was $5.00 (Iowa State University minimum); postage and handling are included.

Charge to user = computer time + postage and handling + network overhead

= $5.00 (Iowa State minimum) + network overhead

CONTENTS—CGELG

pages

1 Identification & Abstract
3,4 User Instructions
5,6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Minres Method of Factor Analysis
CALLING NAME	MINRES
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service (ETS)
AUTHOR(S) AND AFFILIATION(S)	Program due to: H. Harman, ETS W. Jones, National Security Agency Y. Fukuda, TRW Systems, Inc. Adaptation due to: J. Barone, ETS
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Ernest Anastasio, Office of Data Analysis Research, Educational Testing Service, Rosedale Road, Princeton, New Jersey 08540 Tel.: (609) 921-9000 ext. 2552

FUNCTIONAL ABSTRACT

The word "minres" is a contraction of "minimum residuals," and designates a method of factor analysis involving the minimization of off-diagonal residuals of a correlation matrix. Such a method, which has long been sought, has many features that recommend it for initial factorization of a correlation matrix. While the objective of the principal-factor method is to extract maximum variance, the objective of the minres method is to "best" reproduce the observed correlations. The latter objective can be traced to Thurstone, "The object of a factor problem is to account for the tests, or their intercorrelations, in terms of a small number of derived variables, the smallest possible number that is consistent with acceptable residual errors" (Ref. 1, p. 61). This problem has been resolved (Ref. 2, Chap. 9) by minimizing the residual correlations (i.e., the differences between the observed values and those reproduced from the factor-analysis model).

The basic factor-analysis model may be put in the form

1. $z_j = a_{j1}-F_1 + a_{j2}-F_2 + \cdots + a_{jm}-F_m + d_j-U_j \quad (j = 1,2,\cdots,n)$

continued

in which the common factor loadings in the matrix $A = (a_{jp})$ are the only parameters to be estimated. However such a solution is obtained, the matrix $R^{\dagger}$ of reproduced correlations with communalities in the principal diagonal is given by (Ref. 2, p. 28),

2. $R^{\dagger} - AA'$,

where uncorrelated factors are assumed without loss of generality. The condition for a least-squares best fit to the off-diagonal correlations may be expressed as follows.

3. $$f(A) = \sum_{k=j+1}^{n} \sum_{j=1}^{n-1} n \left[r_{jk} - \sum_{p=1}^{m} a_{jp}a_{kp} \right]^2 = \min \quad .$$

The objective function in (3) is to be minimized under the constraints.

4. $$h_j^2 = \sum_{p=1}^{m} a_{jp}^2 \leq 1 \qquad (j=1,2,\cdots,n) \quad .$$

Thus, the object of the minres method is to minimize the function f(A) for a specified number of factors m by varying the values of the factor loadings. The communalities (restricted to numbers between 0 and 1) are obtained as a byproduct of the method.

The basic mathematical method employed in the computer program involves the Gauss-Seidel process whereby successive displacements are introduced in only one row of A at a time, making the objective function f(A) quadratic. The mathematical technique for minimizing a quadratic-functions subject to side conditions involving inequalities is extremely difficult. However, the inequalities can be removed from the problem at hand, so that it can be made tractable (Ref. 2, p. 192).

The computation of factor loadings is continued until the following criterion is met,

5. $$\max_{j,p} \left| {}_{(i)}a_{jp} - {}_{(i-1)}a_{jp} \right| < \varepsilon \qquad (j = 1,\cdots,n;\ \ p = 1,\cdots,m) \ ,$$

when i is the iteration number and ε is preset; or until some maximum number of iterations is reached.

The test statistic, U_m (Ref. 2, p. 197), for testing the significance of m factors is calculated at the conclusion of the program. This is asymptotically dustributed as X^2 with $\frac{1}{2}[(n - m)^2 + n - m]$ degrees of freedom, where n is the number of variables.

continued

The computer output includes: (1) input parameters and correlation matrix; (2) initial factor matrix (m principal components; alternative is a principal-factor solution dependent on some communality input); (3) minres solution, along with derived communalities; (4) value of the objective function f(A) for this solution; (5) matrix of residual correlations; (6) frequency distribution of the residuals; (7) frequency distribution of the differences in factor loadings between the final iteration and the preceding one; (8) test of significance for the number of common factors; and (9) time for all the preceding calculations.

REFERENCES

Thurstone, L.L., *Multiple-Factor Analysis*, (University of Chicago Press, Chicago, Ill., 1947).

Harman, H.H., *Modern Factor Analysis*, Second Edition, Revised, (University of Chicago, Ill., 1967).

USER INSTRUCTIONS

Description of Input

System Control Cards

JOB Card will be prepared by ETS personnel.

System Card 1

Columns	*Contents*
1–36	//STEPNAME EXEC GITNGO,NAME=MINRES

System Card 2

Columns	*Contents*
1–19	//GO.SYSIN DD *

System Card 3

Columns	*Contents*
1– 2	//

Problem Card

Parameter	*Columns*	*Format*	*Description*
NV	1– 5	I5	Number of variables ($\leq$ 66)
M1 M2	6–10 11–15	I5 I5	Range of factors: M1 $\geqq$ 1, M2 $\leqq$ 15
ITMAX	16–20	I5	Maximum number of iterations (usually 1000)
IND	21–25	I5	Initial factor matrix option: 1: Principal factor (dependent upon communality input) 0: Principal components
EPS	26–35	F10.3	Convergence criterion (usually ε = .001)
NPR	36–40	I5	0: Final values printed out
NSAMP	41–45	I5	Number of cases (if blank, test of significance for numbers of factors is omitted)

Title Card

Name of problem in Cols. 1–80.

continued

Format Card

The format specification starts in Col. 2 and may not extend beyond Col. 72. The word FORMAT is *not* used. The standard FORTRAN variable format is followed.

Data Cards

Using specified format, keypunch complete correlation matrix by rows with unities in diagonal, starting each variable on a new card. If initial factor matrix option is used, keypunch communalities on one card (or more as needed), inserting a Format Card immediately before the Communalities Card(s). The communality entries consist of the diagonal entries in the correlation matrix if these are not equal to one.

Order of Cards

JOB Card
System Card 1
System Card 2
Problem Card
Title Card
Format Card
Data (correlation matrix with unities in diagonal)
If initial factor matrix option is used (1 in Col. 25 of Problem Card), insert here
    (a) Format Card
    (b) Data (communality values as desired)
Blank Card
/*
System Card 3

Repeat for additional problems

Output Description

See Functional Abstract for a description of the output of MINRES.

SAMPLE INPUT

```
//            JOB  (....
//STEPNAME  EXEC  GITNGO,NAME=MINRES
//GO.SYSIN  DD     *
     8     2     4 1000     1       .001     0     1

 (8F8.3)
   0.870
   0.846   0.917
   0.805   0.881   0.849
   0.859   0.826   0.801   0.838
   0.473   0.376   0.380   0.436   0.875
   0.398   0.326   0.319   0.329   0.762   0.671
   0.301   0.277   0.237   0.327   0.730   0.583   0.640
   0.382   0.415   0.345   0.365   0.629   0.577   0.539   0.536
 (8F8.3)
   0.870   0.917   0.849   0.838   0.875   0.671   0.640   0.536

/*
//
```

SAMPLE OUTPUT

```
                              2 FACTORS


INITIAL FACTOR MATRIX                MINIMUM RESIDUAL FACTOR MATRIX
                                                 32 ITERATIONS


           1        2    COMMUNALITY           1        2    COMMUNALITY
   VAR.                                VAR.
    1    0.860    0.328    0.846        1    0.856    0.352    0.839
    2    0.850    0.414    0.894        2    0.849    0.412    0.890
    3    0.812    0.412    0.829        3    0.809    0.409    0.821
    4    0.835    0.345    0.815        4    0.831    0.342    0.809
    5    0.746   -0.562    0.872        5    0.750   -0.571    0.889
    6    0.632   -0.499    0.648        6    0.630   -0.492    0.639
    7    0.575   -0.525    0.606        7    0.568   -0.509    0.582
    8    0.611   -0.359    0.502        8    0.607   -0.351    0.491


VARIANCE                             VARIANCE
         4.477    1.537    6.014              4.449    1.510    5.960


                                     OBJECTIVE FUNCTION =    0.0120602772



RESIDUALS

          1        2        3        4        5        6        7        8


1      0.0
2     -0.014    0.0
3     -0.020    0.027    0.0
4      0.037   -0.020   -0.011    0.0
5      0.016   -0.025    0.007    0.008    0.0
6      0.018   -0.006    0.011   -0.027    0.008    0.0
7     -0.021    0.004   -0.014    0.029    0.012   -0.026    0.0
8     -0.024    0.044   -0.002   -0.020   -0.027    0.021    0.015    0.0
```

continued

```
FREQUENCY DISTRIBUTION                      FREQUENCY DISTRIBUTION
     OF RESIDUALS                                OF DELTA A

 CLASS                                       CLASS
 VALUE FREQUENCY                             VALUE FREQUENCY

 0.050       0                              0.0050     0
 0.045       1                              0.0045     0
 0.040       0                              0.0040     0
 0.035       1                              0.0035     0
 0.030       1                              0.0030     0
 0.025       1                              0.0025     0
 0.020       2                              0.0020     0
 0.015       2                              0.0015     0
 0.010       4                              0.0010     0
 0.005       2                              0.0005     0
 0.0         1                              0.0       15
-0.005       1                             -0.0005     1
-0.010       1                             -0.0010     0
-0.015       2                             -0.0015     0
-0.020       4                             -0.0020     0
-0.025       5                             -0.0025     0
-0.030       0                             -0.0030     0
-0.035       0                             -0.0035     0
-0.040       0                             -0.0040     0
-0.045       0                             -0.0045     0
-0.050       0                             -0.0050     0

MEAN       =    0.0000081201                MEAN       =  -0.0000065375
STD. DEV.  =    0.0211291537                STD. DEV.  =   0.0001312492

U = 80.96  WITH   21   DEGREES OF FREEDOM FOR SAMPLE SIZE OF      305

TIME =  1.35 SECONDS
```

continued

000 0087
000 0087

```
                              3 FACTORS

MINIMUM RESIDUAL FACTOR MATRIX          96 ITERATIONS

            1        2        3    COMMUNALITY
  VAR.
    1    0.860   -0.322    0.160     0.868
    2    0.867   -0.432   -0.243     0.998
    3    0.803   -0.396   -0.031     0.803
    4    0.835   -0.340    0.163     0.839
    5    0.750    0.583    0.112     0.915
    6    0.626    0.492   -0.019     0.635
    7    0.565    0.508   -0.001     0.577
    8    0.611    0.362   -0.181     0.537


VARIANCE
         4.480    1.533    0.158     6.171

OBJECTIVE FUNCTION =        0.0045163482

U = 27.31  WITH 15  DEGREES OF FREEDOM FOR SAMPLE SIZE OF   305

TIME =  1.67   SECONDS


                              4 FACTORS

MINIMUM RESIDUAL FACTOR MATRIX          128 ITERATIONS


            1        2        3        4    COMMUNALITY
   VAR.
    1    0.860    0.329    0.026    0.170      0.878
    2    0.863    0.435    0.031   -0.255      0.999
    3    0.802    0.400    0.042   -0.038      0.805
    4    0.839    0.351   -0.136    0.159      0.871
    5    0.740   -0.547    0.013    0.088      0.855
    6    0.643   -0.513    0.256    0.046      0.744
    7    0.583   -0.555   -0.299   -0.071      0.742
    8    0.605   -0.346    0.055   -0.135      0.507
VARIANCE
         4.497    1.570    0.180    0.154      6.401

OBJECTIVE FUNCTION =        0.0006105804

U = 15.78  WITH 10  DEGREES OF FREEDOM  FOR SAMPLE SIZE OF  305

TIME =  2.01   SECONDS
```

COST ESTIMATE

A typical MINRES user will already have an n x n correlation matrix, where n is a number of variables, and will want to try several values for m, the number of factors. Usually the values of m will not exceed n/3. Several sets of data were used in trial runs using 8, 24 and 30 variables. Because the computer time required depends partially on the specific data employed, the cost of a given run cannot be estimated with great precision. Therefore, the table below displays a maximum cost which is greater than any of the costs experienced in any of the trial runs. On rare occasions the user might present data which would be more costly to run than the table indicates; most actual costs will be less than those shown.

MINRES Maximum Cost Table
(Dollars)

Number of Variables, n	Number of factors, m 3	4	5	6	7	8	9	10	11
8	5	5	5	5					
24	5	5	5	10	10	10	10	10	20
30		5	10	20	20	20	20		

Estimated cost to user for the program listed on Sample Input is $5.00.

Charge to user = computer time + postage and handling + network overhead
= $5.00 + $5.00 + network overhead
= $10.00 + network overhead

CONTENTS—MINRES

pages
1– 3 Identification & Abstract
5– 6 User Instructions
7–10 I/O
11 Cost—Contents

DESCRIPTIVE TITLE	Reformat into Either BCD or EBCDIC
CALLING NAMES	EDP C 005 E (from BCD to EBCDIC) EDP C 005 B (from EBCDIC to BCD)
INSTALLATION NAME	Computation Center, The City College of The City University of New York
AUTHOR(S) AND AFFILIATION(S)	A. J. Cohen G. W. Elder Computation Center, The City College of The City University of New York
LANGUAGE	COBOL
COMPUTER	IBM 7040, IBM 360/50 O.S., IBM 360/30 D.O.S.
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	R. M. Lobou, Operations Supervisor, Computation Center, The City College of The City University of New York, New York, N.Y. 10031 Tel.: (212) 621-2374

FUNCTIONAL ABSTRACT

This program is designed to reformat data and to convert decks from BCD to EBCDIC or from EBCDIC to BCD. It can be used on any second- or third-generation equipment that supports COBOL. In addition to conversion, there exists an option to reorganize the data completely. Interspersed gangpunching is also possible. That is, constants —such as dates or program numbers—may be emitted into the output data.

The input data may appear in mixed-character configuration and yet yield decks in the desired character set. That is, each input card can be punched in either BCD or EBCDIC and will be punched as output in the specified punch configuration.

USER INSTRUCTIONS

Internal Control-Card Setup

Card 1—

Columns	
1– 2	contain C$ and identify the first Control Card
3– 5	indicate the number of copies to be made. Data output is in specified card sets, not in file sets. E.g., 003 indicates three copies: on data cards A,B,C,···, the order of the output cards will be AAA,BBB,CCC,···
80	may contain a B for conversion to BCD or an E for conversion to EBCDIC. This column is blank for straight reproduction without character conversion

Card 2— This is the constant card.

Punch desired constants, if any, in the appropriate columns. These constants will be generated in the corresponding columns of the output cards. The columns specified here must be zeros in cards 3–8. For example, to emit the character Z in Col. 67 of the output cards, punch a Z in Col. 67 of this constant card

Cards 3 and 4—These Control Cards are used in pairs to specify the 10's and units positions of the input card from which information will be duplicated into the output cards.

Example 1. To take information from Col. 12 of input and punch into Col. 80 of output, Col. 80 of card 3 should be a 1 (the 10's value of the sending field) and Col. 80 of Control Card 4 should be a 2 (the units value of the sending field)

Example 2. To reproduce the information from Col. 20 of the input in Col. 20 of the output, punch a 2 in Col. 20 of Control Card 3 and a 0 in Col. 20 of Control Card 4

Thus, Control Cards 3 and 4 designate the source columns for all information to be punched in the output.

Cards 5 and 6—These two cards are used in the same manner as cards 3 and 4 described above. The only difference is that the zone punches of the specified columns are stripped before generation into output fields. Thus, an A from

continued

input will appear as a 1 in the output. (The 12-zone punch of the A is stripped, leaving only the numeric portion.)

All unused columns of these two Control Cards—that is, all columns for which no stripping is to be done—are to be filled with zeros

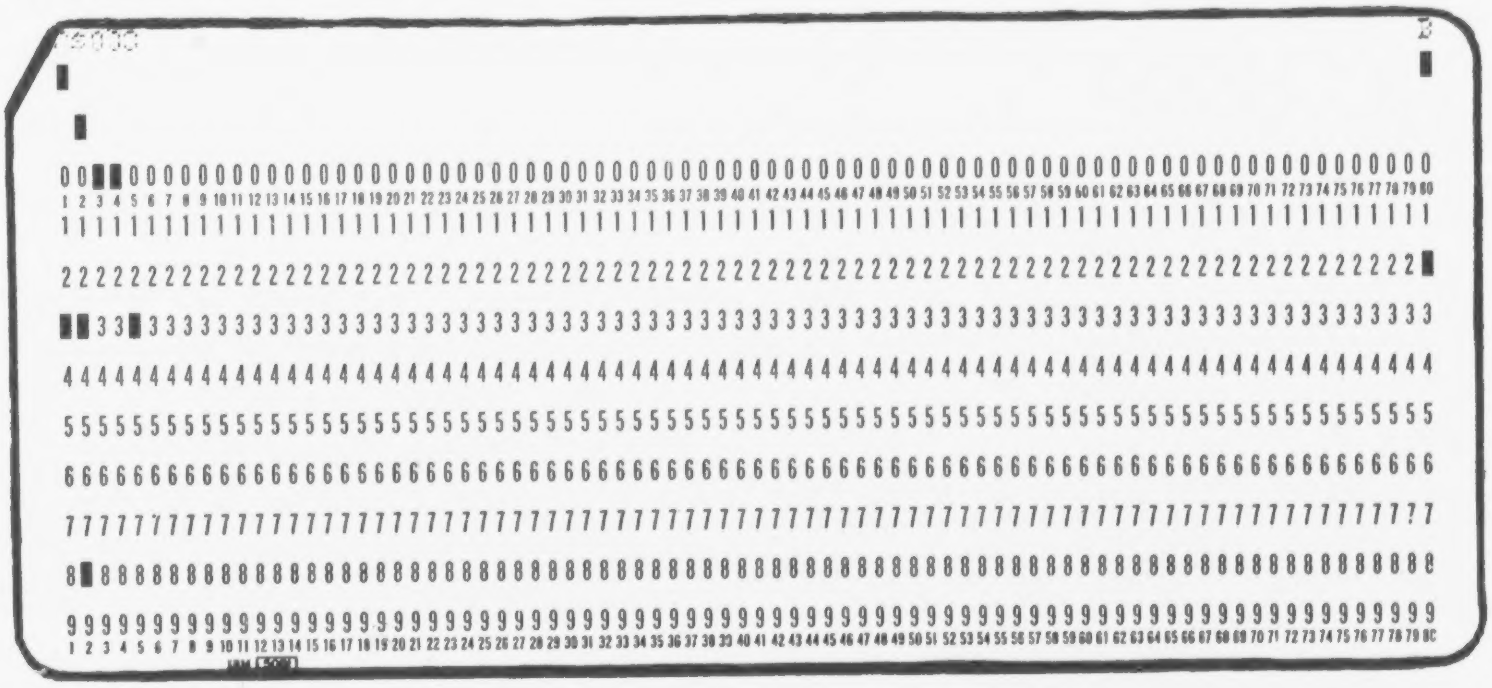
Cards 7 and 8—These two cards perform the same function as Control Cards 5 and 6, except that the numeric portions of the characters are stripped, leaving only the zone punches. Again, the unused fields must be zero-filled

General Instructions

The object module is followed by the eight Control Cards and then by the data to be converted and/or reformatted. Additional Control Cards and data may follow in the job stream without calling the object program again.

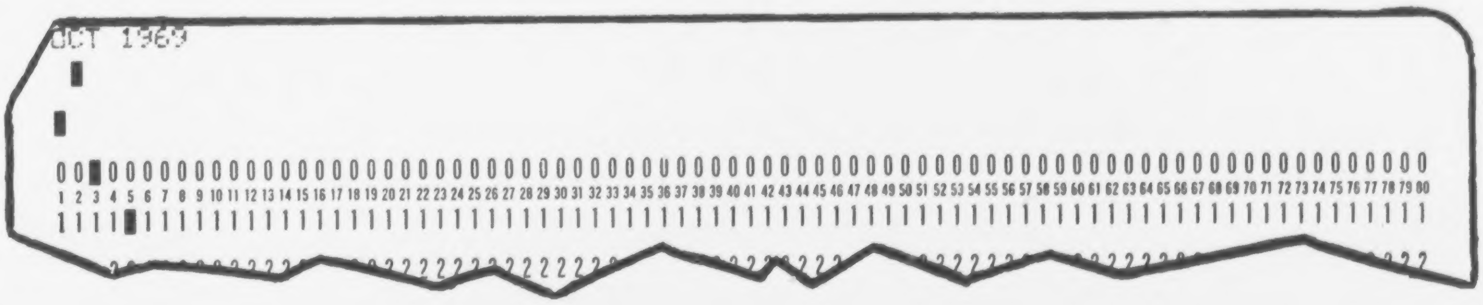
With regard to the eight Control Cards, note that individual column manipulation may be specified in only one set of the three sets of cards (3–4, 5–6, 7–8) or the constant card. Thus, if any manipulation is specified for Cols. 73–80 (either in emitting a constant or specified in one of the pairs of Control Cards), the other Control Cards must be zero in those columns.

SAMPLE INPUT



CARD 1

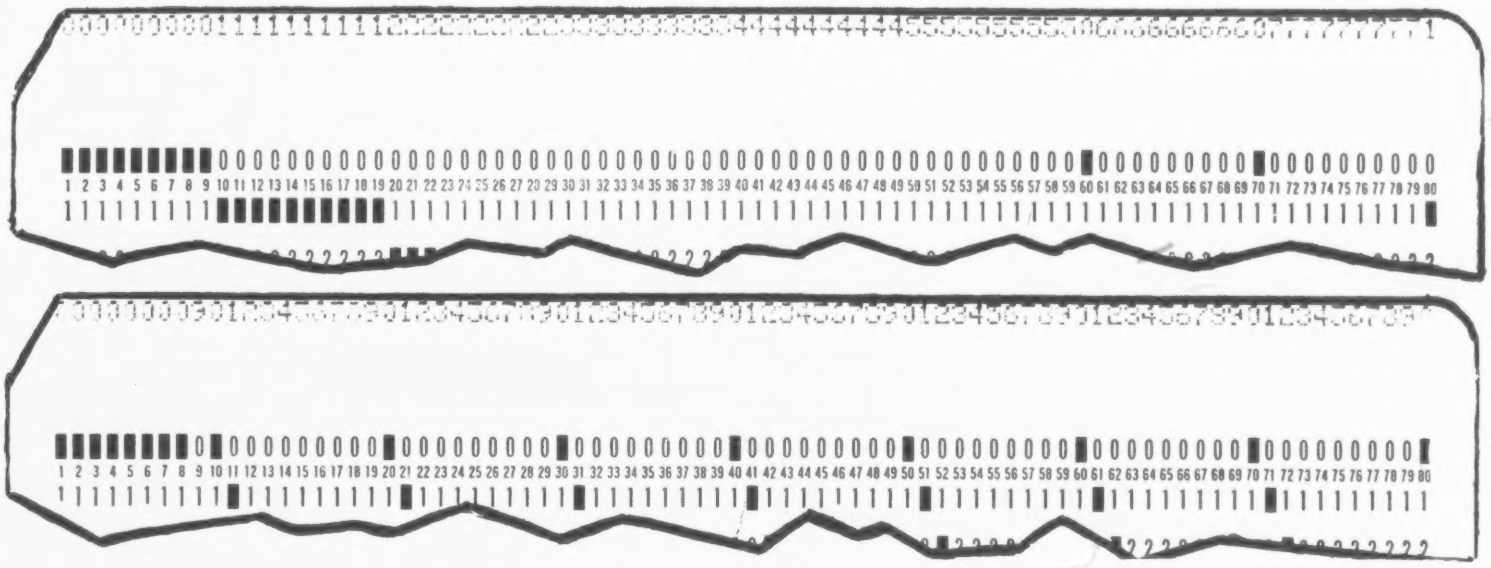
The C$ in Cols. 1–2 indicates the first Control Card. Columns 3–5 specify three copies of output. The B in Col. 80 indicates that the output is to be punched in BCD regardless of the punch configuration of the input.



CARD 2

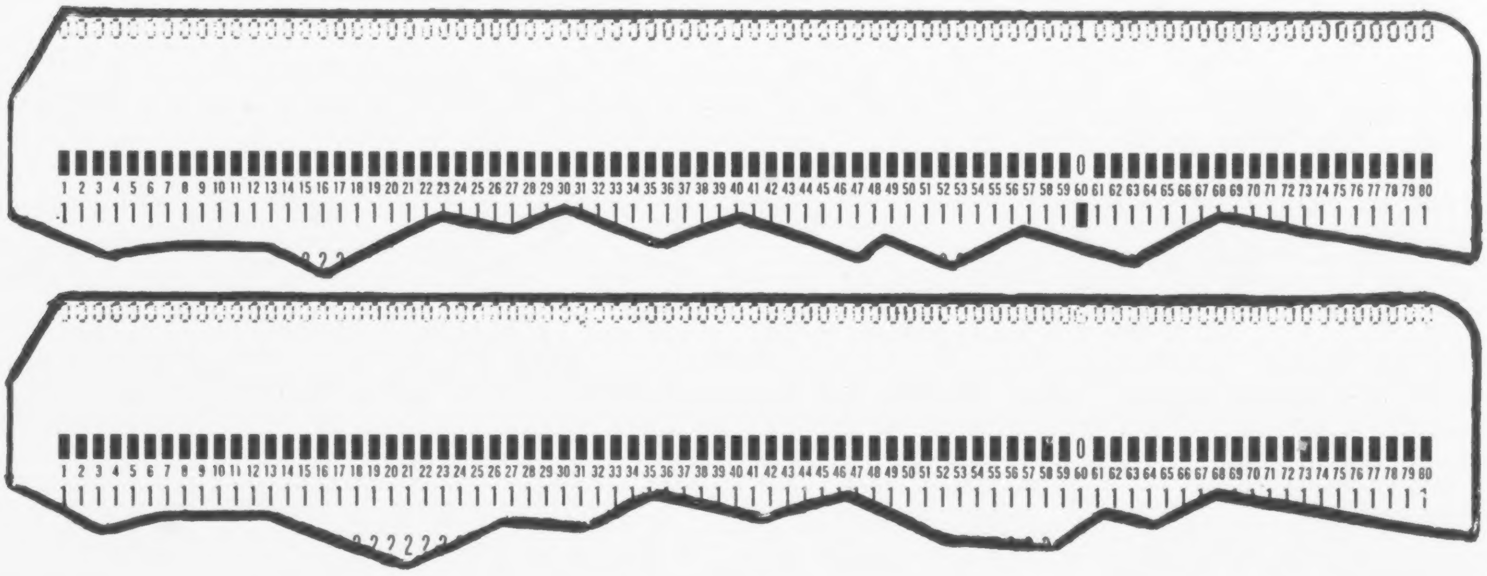
This card specifies that OCT 1969 is to be generated in Cols. 1–8 in each of the output cards.

continued



CARDS 3–4

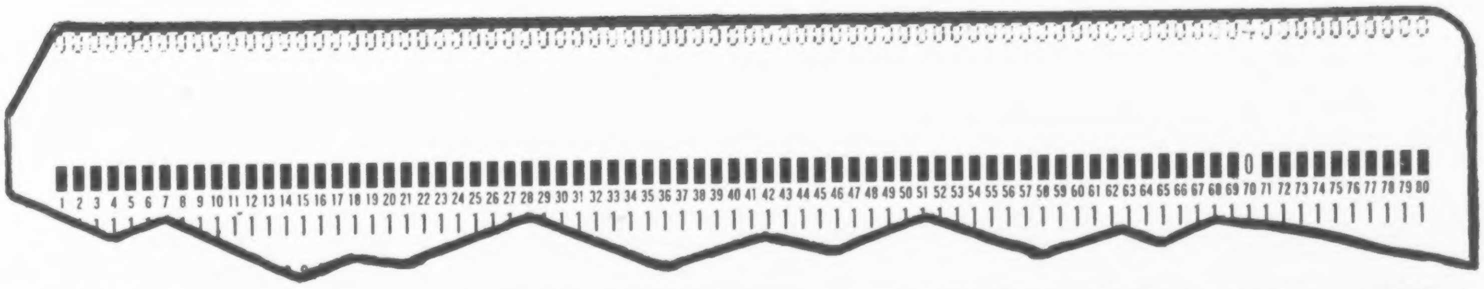
These cards indicate that information from Col. 12 of the input data is to be punched in Col. 80 of output, and all other columns will contain identical data in the input card (unless specified by a subsequent Control Card).

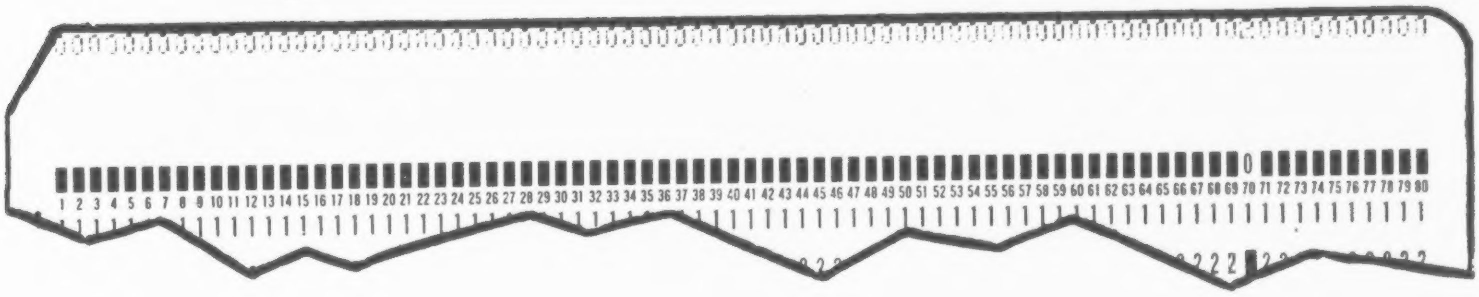


CARDS 5–6

These cards show that the zone portion of the information in Col. 16 is to be stripped before being punched in Col. 60 of the output.

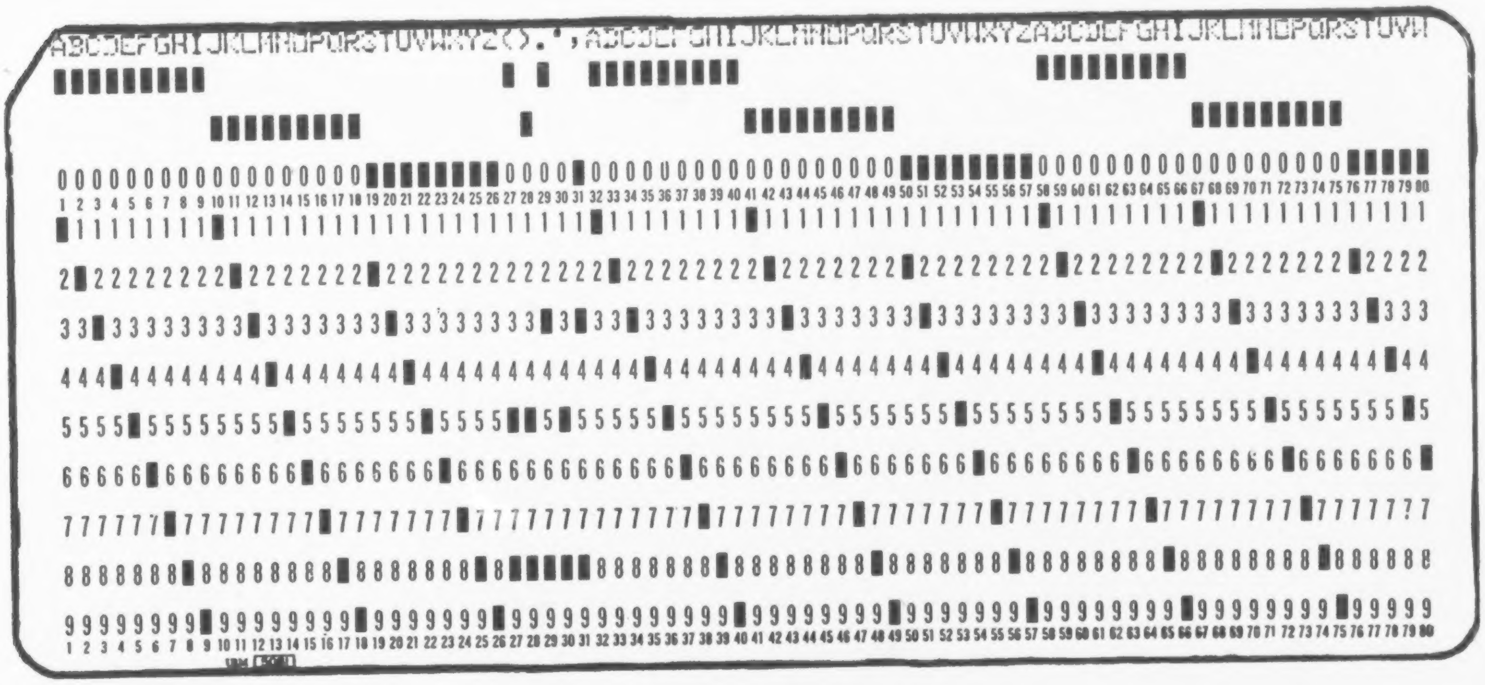
continued





CARDS 7–8

These cards specify that the numeric portion of the information in Col. 42 is to be deleted before being punched in Col. 70 of the output.

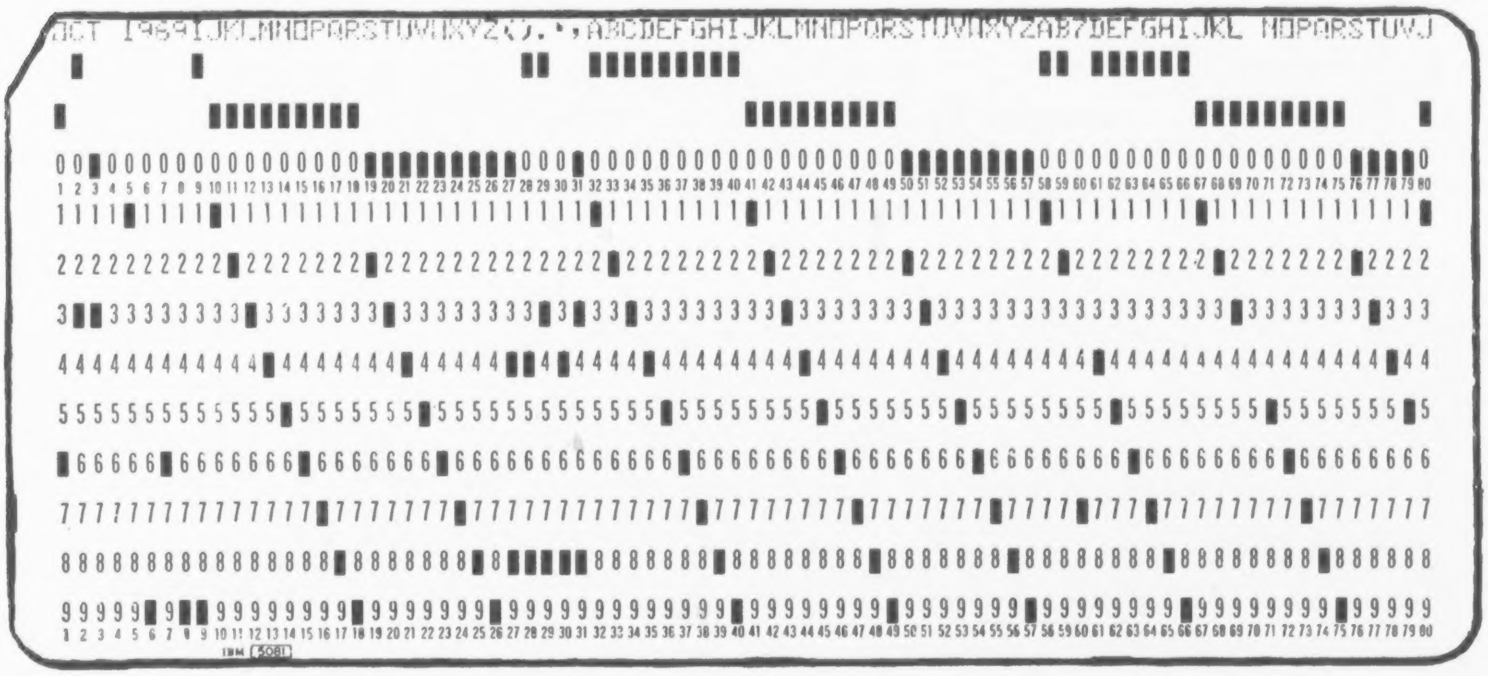

ABCDEFGHIJKLMNOPQRSTUVWXYZ<>.',ABCDEFGHIJKLMNOPQRSTUVWXYZABCDEFGHIJKLMNOPQRSTUVW

SAMPLE DATA CARD

This is one card of a deck of data to be operated upon according to the preceding Control Cards. Note that the card is punched in EBCDIC.

SAMPLE OUTPUT

OCT 1969 has been inserted into the cards. Column 60 contains only the numeric portion of the character in Col. 16 of the original data (see Sample Data Card). Column 70 contains only the zone portion of the character in Col. 70 of the input data; the numeric portion has been deleted. Information from Col. 10 of the original data has been moved to Col. 80 of the output. All other information from the input has been generated in corresponding columns of the output. Note that the output is punched in BCD, as specified in Col. 80 of the first Control Card.


OCT 1969IJKLMNOPQRSTUVWXYZ<().*,ABCDEFGHIJKLMNOPQRSTUVWXYZAB7DEFGHIJKL NOPQRSTUVJ
IBM 5081

(One of three identical cards, as called for in card 1)

continued

COST ESTIMATE

This program will run at standard card read/punch speed where numeric or zone characters are not being eliminated. The speed of the program is reduced, however, as the columnar manipulation is increased.

The cost of a run is computed according to the following formula.

$$\text{COST} = [\text{ET} + (\frac{\text{CR}}{15} + \frac{\text{LP}}{20} + \frac{\text{CP}}{5})/100] \text{ X } \$2.50,$$

where ET is execution time (including CPU time and voluntary wait time), CR is cards read, LP is lines printed, and CP is cards punched.

For example, to reproduce Cols. 1–40 of input data into Cols. 1–40 of output, 1164 cards were read, 41 lines printed and 1046 cards punched, in a total of 0.77 minutes' execution time. Chargeable computer time was $9.20, computed as follows.

$$\begin{aligned}\text{COST} &= [0.77 + (\frac{1164}{15} + \frac{41}{20} + \frac{1046}{5})/100] \text{ X } \$2.50 \\ &= [0.77 + (0.78 + 0.02 + 2.09)/100] \text{ X } \$2.50 \\ &= \$9.20.\end{aligned}$$

Charge to user = computer time + network overhead
= $9.20 + network overhead

80-column reproduction, 1000 cards, requires 0.93 minutes' execution time, and 1-column reproduction, 1000 cards, requires 0.63 minutes' execution time. The average cost is $5.00/1000 cards. Billing will be against a minumum of 5000 cards.

CONTENTS—EDP C 005 E
—EDP C 005 B

pages
1 Identification & Abstract
3,4 User Instructions
5–8 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Analysis of Contingency Tables (ACT version 1.00)
CALLING NAME	G9 *MSU ACT
INSTALLATION NAME	Michigan State University Computer Institute for Social Science Research
AUTHOR(S) AND AFFILIATION(S)	A.M. Lesgold F.M. Sim L.C. Widmayer Computer Institute for Social Science Research Michigan State University
LANGUAGE	FORTRAN and COMPASS
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Anders Johanson, Programming Supervisor, Applications Programming, Computer Laboratory, Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

This program forms bivariate frequency distributions from designated pairs of variables on decks of punched cards or a magnetic tape. Other terms for bivariate frequency distributions include "cross-tabulations", "cross-runs", "two-way breakdowns", "contingency tables", or merely "tables".

The program may also perform any combination of the following operations on designated tables: row and/or column means and standard deviations; percentages of each cell on the associated row, column and/or table totals; theoretical frequencies; cell contributions to table chi-square and degrees of freedom; contingency coefficient; tau; gamma, product-moment correlation coefficient; and Kruskal-Wallis H. Results are printed with appropriate labeling of variables and types of calculations performed.

continued

Gamma and tau will not be computed in any ordinary ACT runs due to the excessive amount of computation time required.

Limitations

A variable field may be as wide as desired, but each field used must be recoded to a maximum of 100 code groups (0,1,...,99). There is no limit on the number of observations (i.e., the number of cards per data deck) which may be used. The maximum number of variables which may be used is 950. These may be on any number of data decks.

USER INSTRUCTIONS

Input Deck

The input to this routine consists of the following eight card types in the order listed.

Comment Cards (optional)
Size Card
Table Description Cards
EOT
Recode Cards (if needed)
Variable Description Cards
EOV
Data Cards

Comment Cards (optional)

The user may punch any information he wishes printed out as a title to his output. It is essential that the first and last cards of this group of cards be blank except for a dollar sign ($) punched in Col. 1 to indicate the beginning and end of the comments. There may be up to 52 comment cards, including the two dollar-sign cards. If the user does not wish to print descriptive information, the Comment Cards (*including* the two-dollar-sign cards) should be omitted from the complete job deck.

Size Card

This card provides information about the user's job. All numbers are right-justified.

Columns	*Contents*
1– 5	The number of observations (subjects, respondents, etc.) in the data set. ACT processes observations until this number or an end-of-file (EOF) occurs. If blank, ACT reads until EOF.
6–10	The number of "raw" variables to be used in the analysis. If the user provides his own READ statement, this field must be zero.
13–15	The number of new "constructed" variables. This includes only those variables (if any) which are constructed in recoding. If no recoding is used, leave these columns blank. If recoding is desired, and

continued

123

Columns	*Contents*
	no new variables are created, punch a 1 in Col. 15. (See Recoding below.) If the user provides his own READ statement, this entry must equal the total number of variables.
16–20	The identification number of the first table.
25	Type of data input 0: Data on cards 1: Data on unblocked tape 2: Data on blocked tape
29–30	Logical unit number of data tape. Must not be logical unit no. 5. For data on cards, punch 60.
35*	Mode of data tape blank: BCD 0: BCD 1: Binary all other entries: Flagged as an error
38–40*	Number of words per observation/logical record. Must be < 300. Only applicable if data are on *blocked* input tape.

Note: Not applicable if data are on cards. (Logical unit 60)

44–45	If data are on cards, the number of cards per observation. If data are on *unblocked* tape, the number of words per observation/logical record.
49–50	Output logical unit number. (Blank means printer.)
51–55	The number of data sets to be processed for this set of control cards.
59	Decode error option: D: ACT will print the observation number and the variable number of the first 25 decode errors for each data set indicated on the Size Card. blank: All decode errors will be dropped by setting them equal to 400000000000000B, which is a negative number, and placed in the undefined category at the bottom of each table.

continued

Columns	*Contents*
60	blank: Print all output 1: Suppress printing of tables
77–80	SIZE

Table Description Cards

Table Description Cards list the variable numbers to be used in the contingency tables and the types of statistics desired in the tables.

There is an arbitrary number of Table Description Cards with arbitrary fields indicating the Calculation Code followed by the variables to be tabled. Each statement on a card is of the form,

L (optional comments) = List.

The optional commments may not contain numbers or an equal sign.

L can take on the values;

C = labels the number which follows to be a Calculations Code,

D = labels the list which follows to be a list of down (control) variables,

A = labels the list which follows to be a list of across (spread) variables,

P = labels the list which follows to be pairs of down and across variables,

P = <control var.> <–> <spread var.> .

A pair is a down variable followed by a hyphen followed by an across variable.

Calculations Codes

The user may specify any combination of the following results to be calculated and printed.

Individual Code	*Calculations*
1	Observed frequencies
2	Table row means and standard deviations
4	Percentages of cells in the table row totals
8	Percentages of cells in table column totals
16	Percentages of cells on table total

continued

Individual Code	*Calculations*
32	Theoretical frequencies
64	Contributions of cells to chi square
128	Table column means and standard deviations
256	Chi square (uncorrected) with degrees of freedom; contingency coefficient
512	Product moment coefficient, tau, gamma

If table row means or table column means are requested, Kruskal-Wallis H will be given automatically.

The code to be punched for desired calculations is determined by adding the codes for each individual calculation. Suppose the user wants an observed frequency distribution (1), with row and column percents (4 and 8), chi square (256), and product-moment correlation (512). The Calculations Code punched is 1 + 4 + 8 + 256 + 512 = 781.

C = 781 or CALC = 781

The Calculations Code must be the first field on the Table Description Cards.

Down and Across List Designators

A list may be a single variable followed by blanks or a comma or a number of variables each followed by commas. A comma does not have to follow the last variable in the list.

A variable may have two forms, a single variable number which is less than 1000 or a number followed by ellipsis marks ... followed by a second number greater than the first; e.g., 72...78 . The second form indicates all numbers X such that $72 \le X \le 78$ are to be used as down or across variables in the tables depending on the list designator which proceeded the list.

For the P designator, the associated list may be composed of one pair or several pairs each followed by a comma.

A down list of several variables implies that *each* variable in the list will be run against *all* the variables in the across list which follows.

EOT Card (follows Table Description Cards)

Columns	*Contents*
1– 3	EOT

Recode Cards (optional)

Variables to be recoded must have a Variable Description Card associated with them (see below). For a variable to be read for recoding, it must be listed on a Table Description Card or on a Recode Card. Each *created* variable mentioned on a Table Description Card must have a Variable Description Card listing only the Variable Number, the Variable Name (optional), and the High Code. A *created* variable is never listed on a Recode Card.

Raw variables to be recoded (and not listed on a Table Description Card) are listed on cards with the following format.

Columns	*Contents*
1– 6	RECODE
13–15	Number of first variable to be recoded (right-justified)
18–20	Second variable
⋮	⋮
78–80	Fourteenth variable.

Repeat with as many Recode Cards as are necessary. Do not skip any fields; the first blank field terminates the reading.

Variable Description Cards

Columns	*Contents*
1– 3	Variable Number
6–37	Variable Name
43–45	The actual card number of the observation for the variable in Cols. 1–3. This has no significance if there is only one card per observation.
49–50	The column number at which the observation for this variable begins.
54–55	The number of columns associated with this variable.
59–60	The high code; must be <99. (A zero high code is set equal to 99.)

Note: If there is more than one card describing a single variable, only the information on the first of these cards

continued

will be used. All subsequent duplications will be printed but their information will be ignored. Therefore, corrections of cards not removed must appear first.

EOV Card (follows Variable Description Cards)

Columns	*Contents*
1– 3	EOV

Data Cards

The data deck is placed here.

If there are several sets of data and/or control cards, the sequence is as follows.

JOB Cards (will be prepared by Applications Programming Group)
Comments Cards (for first tables)
Size Card
Table Description Cards
EOT
Recode Cards
Variable Description Cards
EOV
Data Set 1
End-of-File-Card
Data Set 2
End-of-File
⋮
Comments Cards (for second tables)
⋮

EOF Card (End-of-File)

Columns	*Contents*
1– 2	Multipunch 7 and 8 in both columns.

Recoding

Recoding is accomplished by SUBROUTINE CHANGE, a FORTRAN program in which the user indicates his new codes. For users who are not programmers, the Applications Programming Group will write the program if the user indicates his list of rules for recoding variables or creating new variables.

continued

The general format of such a user-written recoding subroutine is as specified below (FORTRAN; begin in Col. 7).

```
          SUBROUTINE CHANGE(NUB,NVAR)
          COMMON MAT(2)
         *COMMON/B999/INBUF(300)
          IF(NUB)1,1,2
      1   (Initialization only)
          RETURN
      2   (Read and/or recode one observation)
          RETURN
          END
```

*Required only when user provides his own READ statements. *The user must not use the array* INBUF *if ACT reads his data for him.*

SUBROUTINE CHANGE is called once for each data set, for initialization, whether or not recoding is to be done. (NUB=0) If recoding was specified on the Size Card, SUBROUTINE CHANGE will be called once for each observation in the data (NUB=1), to recode each observation. If the user provides his own READ statements, he must signal ACT that he has completed reading by setting NUB=–NUB, followed by a RETURN card.

SUBROUTINE CHANGE cards are inserted into the program and job deck. Job cards will be provided by the Applications Programming Group.

000 0089

SAMPLE INPUT

Control Cards for System CDC 3600

```
'JOB,015525,ACT,3, JOHN TEETER - APPLICATIONS PROGRAMMING
'EQUIP,25=(ACT TEST TAPE), MT(542), RO
'LOAD,25
'RUN,3,3000
```

Program control cards

```
$
     SAMPLE RUN OF ACT FOR EIN SOFTWARE
$
         15          1    0    60                  1          1   D                  SIZE
CALC=1023
D=1          A=5...8
D=2,3        A=6...9
D=13...15      A=4,10...12
EOT
001  VARBL1                                        001    12    01    09
002  VARBL2                                        001    13    02    99
003  VARBL3                                        001    15    02    99
004  VARBL4                                        001    18    01    09
005  VARBL5                                        001    19    02    99
006  VARBL6                                        001    22    01    09
007  VARBL7                                        001    23    02    99
008  VARBL8                                        001    27    02    99
009  VARBL9                                        001    30    01    09
010  VARBL10                                       001    35    02    99
011  VARBL11                                       001    38    01    09
012  VARBL12                                       001    39    02    99
013  VARBL13                                       001    44    01    09
014  VARBL14                                       001    48    01    09
015  VARBL15                                       001    50    02    09
EOV
```

DATA

```
6323A05 01 1 111 176 196 242 020 060 111 213 008 67-1   1
6323A01 02 1 112 246 236 272 020 080 112 227 107 49 3       1
6323A08 03 1 113 216 260 256 020 100 113 225 073 50 3       1
6323A12 04 1 121 238 242 212 040 060 114 242 067 83 2   1
6323A11 05 1 122 234 208 236 040 080 121 337 214 25-3       1
6323A09 06 1 123 238 232 238 040 100 122 391 198 15-1   1
6323A03 07 1 131 294 350 318 060 060 123 385 134 29 0 1
6323A06 08 1 132 308 324 304 060 080 124 373 121 34-3       1
6323A10 09 1 133 320 316 332 060 100 211 364 132 36 1   1
6323A02 10 1 141 215 332 360 100 060 212 351 167 29-4-1-1-1-1
6323A07 11 1 142 282 340 376 100 080 213 382 113 19-2     1
6323A04 12 1 143 220 347 382 100 100 214 399 157 23-3       1
6323B02 13 0 211 164 152 206 020 060 221 250 012 62 2     1
6323B01 14 0 212 166 230 218 020 080 222 207 033 70 1   1
6323B10 15 0 213 240 264 228 020 100 223 229 048 46 0 1
6323B11 16 0 221 210 226 226 040 060 224 211 000 81 4-1-1-1-1
6323B06 17 0 222 230 220 242 040 080 311 306 101 53 1   1
6323B05 18 0 223 218 242 232 040 100 312 281 000 41 0 1
6323B12 19 0 231 336 322 376 060 060 313 264 019 39-1   1
6323B09 20 0 232 376 340 368 060 080 314 299 087 48 1   1
6323B04 21 0 233 322 346 330 060 100 321 312 067 50 2     1
6323B08 22 0 241 274 343 368 100 060 322 301 112 59-1   1
6323B07 23 0 242 321 350 372 100 080 323 279 038 43 0 1
6323B03 24 0 243 307 370 380 100 100 324 286 000 45 2     1
''
```

End-of-file mark

```
THIS IS THE SIZE CARD

 OBS  RAW  NEW  TAB   IN  ILUN  MDE  WLR  PLR OLUN  NDST  DCOD
       15          1   0   60              1          1   D                    SIZE

NO. OF OBSERVATIONS =      99999

NO. OF VARIABLES =            15

AND NEW VARIABLES =           -0

NO RECODING

AVAILABLE SPACE IS        17771  WORDS



THESE ARE THE TABLE DESCRIPTION CARDS

CALC=1023
D=1        A=5,,,8
D=2,3      A=6,,,9
D=13,..15    A=4,10,,,12



THESE ARE THE VARIABLE DESCRIPTION CARDS

      NO,       NAME                                   CRD   COL   FL    HC
      001   VARBL1                                     001   12    01    09
      002   VARBL2                                     001   13    02    99
      003   VARBL3                                     001   15    02    99
      004   VARBL4                                     001   18    01    09
      005   VARBL5                                     001   19    02    99
      006   VARBL6                                     001   22    01    09
      007   VARBL7                                     001   23    02    99
      008   VARBL8                                     001   27    02    99
      009   VARBL9                                     001   30    01    09
      010   VARBL10                                    001   35    02    99
      011   VARBL11                                    001   38    01    09
      012   VARBL12                                    001   39    02    99
      013   VARBL13                                    001   44    01    09
      014   VARBL14                                    001   48    01    09
      015   VARBL15                                    001   50    02    99
      EOV

   PASS      UNUSED CORE (DEC)   DWN   ACR
     1             3584            2     7
     2             5705            3     6
     3             7719            3     7
     4             1249           15     4
     5             6712           15    11
     6            17726           15    12

NUMBER OF TABLES TO BE PROCESSED IS          24


***THE DECODE OPTION IS IN EFFECT***



   THE FOLLOWING IS THE FIRST OBSERVATION FOR THE RAW VARIABLES

  1    2    3    4    5    6    7    8    9   10   11   12   13   14   15
  1    1   11    1   76    1   96   42    0   60    1   11    3    8   67

REVISED NO, OF OBSERVATIONS =        24

SETUP TIME FOR THE FOLLOWING          6  TABLES   0.437  SECONDS
```

continued

```
DATE    04/03/70          SAMPLE RUN OF ACT FOR EIN SOFTWARE                                        VERSION   (2.10)   AC
        TABLE      1        THE DOWN VARIABLE (  1) IS VARBL1   AND THE ACROSS VARIABLE (  5) IS VARBL5

**** VARBL1    =  0

VARBL5  =      0      1      2      3      4      5      6      7      8      9     10     11     12     13     14

  FREQ         0      0      0      0      0      0      0      1      0      0      1      0      0      0      0
  % ACR     0,00   0,00   0,00   0,00   0,00   0,00   0,00   8,33   0,00   0,00   8,33   0,00   0,00   0,00   0,00
  % DWN     0,00   0,00   0,00   0,00   0,00   0,00   0,00 100,00   0,00   0,00 100,00   0,00   0,00   0,00   0,00
  % TOT     0,00   0,00   0,00   0,00   0,00   0,00   0,00   4,16   0,00   0,00   4,16   0,00   0,00   0,00   0,00
  TH FR        0      0      0      0      0      0      0      1      1      0      1      0      0      0      0
  CH SQ     0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,50   0,50   0,00   0,50   0,00   0,00   0,00   0,00

              15     16     17     18     19     20     21     22     23     24     25     26     27     28     29

  FREQ         0      0      0      1      0      0      1      1      0      0      0      0      0      0      0
  % ACR     0,00   0,00   0,00   8,33   0,00   0,00   8,33   8,33   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  % DWN     0,00   0,00   0,00 100,00   0,00   0,00 100,00 100,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  % TOT     0,00   0,00   0,00   4,16   0,00   0,00   4,16   4,16   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  TH FR        1      1      0      1      0      1      1      1      0      0      0      0      0      0      0
  CH SQ     0,50   0,50   0,00   0,50   0,00   1,00   0,50   0,50   0,00   0,00   0,00   0,00   0,00   0,00   0,00

              30     31     32     33     34     35     36     37     38     39     40     41     42     43     44

  FREQ         1      0      0      0      0      0      1      0      0      0      1      0      0      0      0
  % ACR     8,33   0,00   0,00   0,00   0,00   0,00   8,33   0,00   0,00   0,00   8,33   0,00   0,00   0,00   0,00
  % DWN   100,00   0,00   0,00   0,00   0,00   0,00 100,00   0,00   0,00   0,00 100,00   0,00   0,00   0,00   0,00
  % TOT     4,16   0,00   0,00   0,00   0,00   0,00   4,16   0,00   0,00   0,00   4,16   0,00   0,00   0,00   0,00
  TH FR        1      0      0      0      1      0      1      0      1      0      1      0      0      0      0
  CH SQ     0,50   0,00   0,00   0,00   0,50   0,00   0,50   0,00   1,00   0,00   0,50   0,00   0,00   0,00   0,00

              45     46     47     48     49     50     51     52     53     54     55     56     57     58     59

  FREQ         0      0      0      0      0      0      0      0      0      0      0      0      0      0      0
  % ACR     0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  % DWN     0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  % TOT     0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  TH FR        0      1      0      0      0      0      0      0      0      0      0      0      0      0      0
  CH SQ     0,00   0,50   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00

              60     61     62     63     64     65     66     67     68     69     70     71     72     73     74

  FREQ         0      0      0      0      1      0      1      0      0      0      0      0      0      0      1
  % ACR     0,00   0,00   0,00   0,00   8,33   0,00   8,33   0,00   0,00   0,00   0,00   0,00   0,00   0,00   8,33
  % DWN     0,00   0,00   0,00   0,00 100,00   0,00 100,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00 100,00
  % TOT     0,00   0,00   0,00   0,00   4,16   0,00   4,16   0,00   0,00   0,00   0,00   0,00   0,00   0,00   4,16
  TH FR        0      0      0      0      1      0      1      0      0      0      0      0      0      0      1
  CH SQ     0,00   0,00   0,00   0,00   0,50   0,00   0,50   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,50

              75     76     77     78     79     80     81     82     83     84     85     86     87     88     89

  FREQ         0      1      0      0      0      0      0      0      0      0      0      0      0      0      0
  % ACR     0,00   8,33   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  % DWN     0,00  50,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  % TOT     0,00   4,16   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,00
  TH FR        0      1      0      0      0      0      0      1      0      0      0      0      0      0      0
  CH SQ     0,00   0,00   0,00   0,00   0,00   0,00   0,00   0,50   0,00   0,00   0,00   0,00   0,00   0,00   0,00
```

continued

000 0089

```
              90      91      92      93      94
  FREQ         0       0       0       0       0
  % ACR     0,00    0,00    0,00    0,00    0,00                                           TOTAL    MEAN   ST DV
  % DWN     0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    0,00    0,00    0,00    0,00                                              12   38,66   25,14
  TH FR        0       0       0       0       1                                          100,00
  CH SQ     0,00    0,00    0,00    0,00    0,50                                           50,00
                                                                                           50,00
**** VARBL1    * 1

VARBL5  =      0       1       2       3       4       5       6       7       8       9      10      11      12      13      14

  FREQ         0       0       0       0       0       0       0       0       1       0       0       0       0       0       0
  % ACR     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    8,33    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00
  TH FR        0       0       0       0       0       0       0       1       1       0       1       0       0       0       0
  CH SQ     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,50    0,50    0,00    0,50    0,00    0,00    0,00    0,00

              15      16      17      18      19      20      21      22      23      24      25      26      27      28      29

  FREQ         1       1       0       0       0       2       0       0       0       0       0       0       0       0       0
  % ACR     8,33    8,33    0,00    0,00    0,00   16,66    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN   100,00  100,00    0,00    0,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     4,16    4,16    0,00    0,00    0,00    8,33    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  TH FR        1       1       0       1       0       1       1       1       0       0       0       0       0       0       0
  CH SQ     0,50    0,50    0,00    0,50    0,00    1,00    0,50    0,50    0,00    0,00    0,00    0,00    0,00    0,00    0,00

              30      31      32      33      34      35      36      37      38      39      40      41      42      43      44

  FREQ         0       0       0       0       1       0       0       0       2       0       0       0       0       0       0
  % ACR     0,00    0,00    0,00    0,00    8,33    0,00    0,00    0,00   16,66    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN     0,00    0,00    0,00    0,00  100,00    0,00    0,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    0,00    0,00    0,00    4,16    0,00    0,00    0,00    8,33    0,00    0,00    0,00    0,00    0,00    0,00
  TH FR        1       0       0       0       1       0       1       0       1       0       1       0       0       0       0
  CH SQ     0,50    0,00    0,00    0,00    0,50    0,00    0,50    0,00    1,00    0,00    0,50    0,00    0,00    0,00    0,00

              45      46      47      48      49      50      51      52      53      54      55      56      57      58      59

  FREQ         0       1       0       0       0       0       0       0       0       0       0       0       0       0       0
  % ACR     0,00    8,33    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN     0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  TH FR        0       1       0       0       0       0       0       0       0       0       0       0       0       0       0
  CH SQ     0,00    0,50    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00

              60      61      62      63      64      65      66      67      68      69      70      71      72      73      74

  FREQ         0       0       0       0       0       0       0       0       0       0       0       0       0       0       0
  % ACR     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  TH FR        0       0       0       0       1       0       1       0       0       0       0       0       0       0       1
  CH SQ     0,00    0,00    0,00    0,00    0,50    0,00    0,50    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,50

              75      76      77      78      79      80      81      82      83      84      85      86      87      88      89

  FREQ         0       1       0       0       0       0       0       1       0       0       0       0       0       0       0
  % ACR     0,00    8,33    0,00    0,00    0,00    0,00    0,00    8,53    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN     0,00   50,00    0,00    0,00    0,00    0,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    4,16    0,00    0,00    0,00    0,00    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  TH FR        0       1       0       0       0       0       0       1       0       0       0       0       0       0       0
  CH SQ     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,50    0,00    0,00    0,00    0,00    0,00    0,00    0,00
```

continued

```
             90      91      92      93      94

  FREQ        0       0       0       0       1                                TOTAL   MEAN  ST DV
  % ACR    0,00    0,00    0,00    0,00    8,33
  % DWN    0,00    0,00    0,00    0,00  100,00                                   12  40,58  28,78
  % TOT    0,00    0,00    0,00    0,00    4,16                               100,00
  TH FR       0       0       0       0       1                                50,00
  CH SQ    0,00    0,00    0,00    0,00    0,50                                50,00

**** T O T A L S ****

VARBL5 =      0       1       2       3       4       5       6       7       8       9      10      11      12      13      14

  FREQ        0       0       0       0       0       0       0       1       1       0       1       0       0       0       0
  MEAN     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR    0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16    4,16    0,00    4,16    0,00    0,00    0,00    0,00
  % DWN    0,00    0,00    0,00    0,00    0,00    0,00    0,00  100,00  100,00    0,00  100,00    0,00    0,00    0,00    0,00
  % TOT    0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16    4,16    0,00    4,16    0,00    0,00    0,00    0,00

             15      16      17      18      19      20      21      22      23      24      25      26      27      28      29

  FREQ        1       1       0       1       0       2       1       1       0       0       0       0       0       0       0
  MEAN     1,00    1,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR    4,16    4,16    0,00    4,16    0,00    8,33    4,16    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN  100,00  100,00    0,00  100,00    0,00  100,00  100,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT    4,16    4,16    0,00    4,16    0,00    8,33    4,16    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00

             30      31      32      33      34      35      36      37      38      39      40      41      42      43      44

  FREQ        1       0       0       0       1       0       1       0       2       0       1       0       0       0       0
  MEAN     0,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR    4,16    0,00    0,00    0,00    4,16    0,00    4,16    0,00    8,33    0,00    4,16    0,00    0,00    0,00    0,00
  % DWN  100,00    0,00    0,00    0,00  100,00    0,00  100,00    0,00  100,00    0,00  100,00    0,00    0,00    0,00    0,00
  % TOT    4,16    0,00    0,00    0,00    4,16    0,00    4,16    0,00    8,33    0,00    4,16    0,00    0,00    0,00    0,00

             45      46      47      48      49      50      51      52      53      54      55      56      57      58      59

  FREQ        0       1       0       0       0       0       0       0       0       0       0       0       0       0       0
  MEAN     0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00

             60      61      62      63      64      65      66      67      68      69      70      71      72      73      74

  FREQ        0       0       0       0       1       0       1       0       0       0       0       0       0       0       1
  MEAN     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR    0,00    0,00    0,00    0,00    4,16    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16
  % DWN    0,00    0,00    0,00    0,00  100,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00  100,00
  % TOT    0,00    0,00    0,00    0,00    4,16    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16

             75      76      77      78      79      80      81      82      83      84      85      86      87      88      89

  FREQ        0       2       0       0       0       0       0       1       0       0       0       0       0       0       0
  MEAN     0,00    0,50    0,00    0,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV    0,00    0,70    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR    0,00    8,33    0,00    0,00    0,00    0,00    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN    0,00  100,00    0,00    0,00    0,00    0,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT    0,00    8,33    0,00    0,00    0,00    0,00    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00
```

continued

134

FILMED FROM BEST AVAILABLE COPY

```
VARBL5  =      0       1       2       3       4       5       6       7       8       9      10      11      12      13      14

  FREQ         0       0       0       0       0       0       0       1       1       0       1       0       0       0       0
  MEAN      0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR     0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16    4,16    0,00    4,16    0,00    0,00    0,00    0,00
  % DWN     0,00    0,00    0,00    0,00    0,00    0,00    0,00  100,00  100,00    0,00  100,00    0,00    0,00    0,00    0,00
  % TOT     0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16    4,16    0,00    4,16    0,00    0,00    0,00    0,00

              15      16      17      18      19      20      21      22      23      24      25      26      27      28      29

  FREQ         1       1       0       1       0       2       1       1       0       0       0       0       0       0       0
  MEAN      1,00    1,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR     4,16    4,16    0,00    4,16    0,00    8,33    4,16    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN   100,00  100,00    0,00  100,00    0,00  100,00  100,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     4,16    4,16    0,00    4,16    0,00    8,33    4,16    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00

              30      31      32      33      34      35      36      37      38      39      40      41      42      43      44

  FREQ         1       0       0       0       1       0       1       0       2       0       1       0       0       0       0
  MEAN      0,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR     4,16    0,00    0,00    0,00    4,16    0,00    4,16    0,00    8,33    0,00    4,16    0,00    0,00    0,00    0,00
  % DWN   100,00    0,00    0,00    0,00  100,00    0,00  100,00    0,00  100,00    0,00  100,00    0,00    0,00    0,00    0,00
  % TOT     4,16    0,00    0,00    0,00    4,16    0,00    4,16    0,00    8,33    0,00    4,16    0,00    0,00    0,00    0,00

              45      46      47      48      49      50      51      52      53      54      55      56      57      58      59

  FREQ         0       1       0       0       0       0       0       0       0       0       0       0       0       0       0
  MEAN      0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR     0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN     0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00

              60      61      62      63      64      65      66      67      68      69      70      71      72      73      74

  FREQ         0       0       0       0       1       0       1       0       0       0       0       0       0       0       1
  MEAN      0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV     0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR     0,00    0,00    0,00    0,00    4,16    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16
  % DWN     0,00    0,00    0,00    0,00  100,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00  100,00
  % TOT     0,00    0,00    0,00    0,00    4,16    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00    4,16

              75      76      77      78      79      80      81      82      83      84      85      86      87      88      89

  FREQ         0       2       0       0       0       0       0       1       0       0       0       0       0       0       0
  MEAN      0,00    0,50    0,00    0,00    0,00    0,00    0,00    1,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  ST DV     0,00    0,70    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % ACR     0,00    8,33    0,00    0,00    0,00    0,00    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % DWN     0,00  100,00    0,00    0,00    0,00    0,00    0,00  100,00    0,00    0,00    0,00    0,00    0,00    0,00    0,00
  % TOT     0,00    8,33    0,00    0,00    0,00    0,00    0,00    4,16    0,00    0,00    0,00    0,00    0,00    0,00    0,00
```

COST ESTIMATE

Approximate cost is 25¢ per table when the number of observations is less than 5000, slightly higher when it is above. In addition, there is a consulting fee of $10./hr. for work done by the Applications Programming Group.

Approximate charge to user = computer time + consulting + network overhead

CONTENTS—G9 *MSU ACT

pages
1– 2 Identification & Abstract
3– 9 User Instructions
11–16 I/O
17 Cost—Contents

DESCRIPTIVE TITLE	Student Scheduling System
CALLING NAME	SCHEDULE
INSTALLATION NAME	The Michigan State University Computer Laboratory, Applications Programming
AUTHOR(S) AND AFFILIATION(S)	Withheld by request
LANGUAGE	FORTRAN IV and COMPASS
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Available for service at Michigan State University only
CONTACT	Mr. Anders Johanson, 324 Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355–4684

FUNCTIONAL ABSTRACT

School scheduling has two major parts, master schedule building and sectioning. Master schedule building is the process of deciding the courses to be offered and their time, place, instructor and seating capacity. Sectioning is the assigning of students into the courses of the master schedule.

The Student Scheduling System provides the school with a few valuable aids for the master schedule building process, performs the sectioning process and prints the individual students' schedules and the class lists.

USER INSTRUCTIONS

Detailed instructions and some samples of input and output are available from the references cited below. In addition, the Programming Applications Group recommends a personal conference with one of its staff members for most effective understanding and use of the system.

Options

The options listed in the instructions are described here to give a more complete picture of the system.

Semesters

The school has the option of either one or two semesters per school year.

Traditional or Flexible Scheduling

Traditional scheduling is requested when all courses meet at exactly the same period every day and on all days. Flexible scheduling is used when any course meets at different times of the day or not on every day of the week or both. The maximum number of modules permitted per day is twenty-nine. (Module is synonymous with period in a traditional schedule.)

Required or Elective Courses

The school also has the option to designate courses as required, necessary for graduation, or elective.

Class and Section Closing

Class section cards furnish the seating capacities of the sections. A section is closed when its seating capacity is reached. A class is closed when all of its sections are closed. For all multi-section courses, an attempt is made to assign seats to all sections evenly. With the special options of unlimited seating capacities and temporarily closing the highly assigned sections, this even distribution is obtained.

Free Periods or Study Halls

Periods left open after all course assignments for a student are called free periods. The school may assign section numbers for study halls, with seating capacities. The school may elect not to assign study halls, in which case a count is given, by period, of the number of students who have free periods. An option is also available to limit the number of consecutive free periods or study halls.

continued

Common Courses

One type of common course is a course that all students must request. The computer will automatically add these common courses to the individual student requests. Another type of common course is one that must be requested by all students of a given grade level. This option is often useful in scheduling lunch or homeroom.

Equivalent and Alternate Courses

Each student, when completing the student request form, can choose an *alternate* course to be used in place of an *elective* that cannot be scheduled. *Equivalent* courses are designated by the school as being similar in major aspects and may be substituted for a requested course (required or elective) that cannot be scheduled.

Course Substitution or Deletion

The "required" designation has little bearing as long as a student can be scheduled for all courses he has requested. However, if conflicts develop, the student is assigned to his required courses before being assigned to any of his electives. If a student has a course conflict and no substitution is possible, then the course in conflict is deleted from his request and his remaining courses are scheduled. This gives the student a partial schedule and reflects his scheduled courses in the overall load.

Operation

The following step-by-step description of the automated sectioning is presented here to give an idea of what the overall process involves.

Initial Information

1. The school administration determines the courses to be offered in the curriculum and assigns a three digit number to each course.
2. The students are counseled regarding the available courses and their requests are recorded on student request cards.
3. The list of courses to be offered and the student request cards are forwarded to M.S.U. and the preliminary phase is run.

continued

The Tally Sheet, Two- and Three-Way Conflict Matrices

1. The list of student request cards indicates errors and courses not offered.

2. The tally sheet provides totals of students requesting each course by grade level code, an aid in determining the number of sections that should be offered for each course.

3. The two-way conflict matrix provides the school with potential conflicts in the schedule for particular course combinations. The three-way conflict matrix shows conflicts existing between a given course and a course pair. The matrices are printed in table form in course-number order.

4. A no-conflict listing is also provided with listings of courses that have no conflicts; i.e., all students have requested only one of the courses from each list.

Master Schedule

1. The master schedule is now created by the school administrators using the tally sheet and the conflict matrices. Sections, seating, instructors, rooms and times are indicated.

2. The master schedule is submitted to the Computer Center for a trial run along with any changes in the students' requests.

Output of the Trial Runs

1. The output of the trial run includes an updated listing of the current school file, a list of rejected student requests with the reason for rejection, the master schedule with number of seats assigned and number of seats remaining for each section, a list of closed sections, and a final summary that includes the number of students scheduled and number rejected.

2. The first trial run is normally without the substitution of alternate or equivalent courses. In this first run, students whose courses cannot be scheduled completely are given a partial schedule in order to reflect their incomplete schedule in the overall load. This output information is now used to modify the master schedule where necessary.

3. Corrections and modifications to the master schedule are submitted along with any student request changes.

continued

4. In the second trial run, alternate and equivalent courses will be substituted where required. This run produces additional output consisting of a list of students with partial schedules. These lists are used by the school to modify student requests. If, however, the school finds it necessary to change the master schedule at this time, then the second trial run should be re-run using the latest master schedule.

5. The final master schedule and any necessary student request changes are then submitted for the final run.

Final Run

1. The output of the final run includes a master schedule with the number of seats assigned and number remaining in each section, a list of students who had course changes with alternate and/or equivalent courses, a list of rejected students with partial schedules and courses causing conflict, multi-copies of an individual schedule for each student, and a class list for each section.

The total length of time required to complete the student scheduling process is primarily dependent upon the school. The time between our receiving the information from the school and shipping the output to the school is normally less than one week for each step. Therefore, the sooner the Computer Laboratory receives the correct information for each step of the scheduling process, the sooner the school receives a completed set of output from the automated sectioning system.

We suggest that the school complete at least the first trial run before the end of the school year so that the administrators can counsel both those students affected by any change in the master schedule and those students with problems generated from their own requested courses.

REFERENCES

Michigan State University Computer Laboratory, Applications Programming Group, *Student Scheduling System: Description* (East Lansing, Mich., 1969).

Michigan State University Computer Laboratory, Applications Programming Group, *Student Scheduling System: User Manual* (East Lansing, Mich., 1968).

Both these references are on file with the Educational Information Network, and copies are available at the cost of reproduction.

SAMPLE INPUT AND SAMPLE OUTPUT

Input/Output samples are included in the references cited in the User Instructions. Persons desiring to obtain a copy of these documents should contact the EIN office.

COST ESTIMATE

Services included in use of SCHEDULE are a preliminary run, two trial runs, a final run and all the output from these runs with three copies of student schedules and two copies of each class section. Charges for these services are as follows.

$.60/student	for >500 students
$.65/student	for <500 students

Each additional trial run is charged $10.00 + $.01/student.

Charge to user = computer costs + network overhead

CONTENTS—SCHEDULE

pages

1 Identification & Abstract
3– 6 User Instructions
7 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Matrix Operations
CALLING NAME	AES 106
INSTALLATION NAME	The Michigan State University Computer Laboratory, Applications Programming
AUTHOR(S) AND AFFILIATION(S)	William L. Ruble Richard J. Martz Agricultural Experimental Station Michigan State University
LANGUAGE	3600 COMPASS (closed subroutines callable from FORTRAN 600 or COMPASS)
COMPUTER	CDC 3600
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Anders Johanson, 324 Computer Center, Michigan State University, East Lansing, Mich. 48823 Tel.: (517) 355-4684

FUNCTIONAL ABSTRACT

This is a package of several subroutines for double-precision matrix manipulation.

REFERENCES

Michigan State University Agricultural Experimental Station, *AES MISC Series Description Numbers 106C, 106D, 106E, 106F, 106I, 106J, 106K, 106L.* Mimeographed descriptions available through Mr. Johanson, above, or from EIN at the cost of reproduction.

USER INSTRUCTIONS

A brief description of each of the subroutines' capabilities is given below. Further details and the actual subroutines are available through the contact person. All operands and results are in double-precision form.

106C - DMADD

This subroutine will add, subtract, element-multiply, or element-divide two matrices (or double or square the elements of a matrix), all matrices being stored columnwise. The dimensions of the matrices are specified as arguments, hence may be varied.

106D - DMLINC1

This subroutine will multiply a matrix by a scalar and add it to another matrix or multiply two matrices by scalars and add them together to form a new matrix. All matrices are stored columnwise. The original matrices may remain unchanged, or the new matrix may be stored over one of the original matrices. The row dimensions of the matrices are specified as arguments, hence may be varied.

106E - DMMUL

This subroutine will multiply two matrices, square a matrix, pre- or post-multiply a matrix by the transpose of another matrix, or pre- or post-multiply a matrix by its own transpose.

106F - DMVM

This subroutine will multiply a matrix or the transpose of a matrix by a vector.

106I - DSCALEM

This subroutine will multiply a matrix by a scalar. The resulting matrix may be stored as a new matrix or over the original matrix. The row dimensions of the matrices are specified as arguments, hence may be varied.

106J - DVINNER

This subroutine will obtain the inner product of two vectors (the sum of the cross-products of the vectors), or obtain the sum of squares of a vector.

continued

106K - DVVT

This subroutine will post-multiply a column vector by a row vector to form a matrix. The dimension of the matrix is specified as an argument, hence can be varied.

106L - TRIDIN

This subroutine will invert a symmetric, positive definite matrix (stored columnwise, upper triangular) and provide solutions to the matrix question AX = B where A and B are given matrices.

REFERENCES

Michigan State University Agricultural Experimental Station, *AES MISC Series Description Number 106C, 106D, 106E, 106F, 106I, 106J, 106K, 106L.* Mimeographed descriptions available through Mr. Johanson, above, or from EIN at the cost of reproduction.

SAMPLE INPUT AND OUTPUT

Sample jobs for the routines described are not presented here, since these routines must be embedded within a larger program that handles actual data input and output.

COST ESTIMATE

Since the subroutines described above can only be used within the context of a more extensive program, cost estimates for the subroutines themselves would not be very helpful. Persons seeking specific information are directed to the contact person.

Charge to user = computer costs + network overhead

CONTENTS—AES 106

pages
1 Identification & Abstract
3– 4 User Instructions
5 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Mineral Identification
CALLING NAME	MINERAL
INSTALLATION NAME	Kiewit Computation Center Dartmouth College
AUTHOR(S) AND AFFILIATION(S)	Michael Hebb Kiewit Computation Center Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

MINERAL was written as an aid in mineral identification. It may be used in two ways. As a reference tool, it will print all stored data on any of 601 different minerals. As an aid in mineral identification it will sort through all stored minerals and list only those that fit the data supplied by the user. The program requests data, sorts, and prints the names of the minerals having similar properties (up to 25).

The scan routine for refractive-index matches has a tolerance of ± .005; hardness has a tolerance of ± .51; and specific gravity has a tolerance of ± .21.

The user has the option to supply more data on the same mineral if he is not satisfied with the preliminary match Several minerals may be checked or referenced in a single run.

REFERENCES

Deer, W.A., Howie, R.A., and Zussman, Jr., *Rock Forming Minerals*, (John Wiley & Sons, Inc., New York, 1962), 5 Vol.

Larsen, E., and Berman, H., *Microscopic Determination of the Nonopaque Minerals*, (U.S. Govt. Print. Off., Wash. D.C., 1934).

USER INSTRUCTIONS

Call the program following the procedures outlined in Ref. 1. Respond 'YES' to the question 'Would you like directions?' Be sure to type all alphabetic input in capital letters.

REFERENCES

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Computation Center, Dartmouth College, Hanover, N.H., June 1970). Available from the EIN Office for the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT .

```
MINERAL    22 SEP 70  14:45

WOULD YOU LIKE DIRECTIONS? YES

     THIS PROGRAM WAS WRITTEN TO AID IN MINERAL IDENT-
IFICATION.  IT MAY BE USED IN TWO WAYS.  AS A REFERENCE
TOOL IT WILL PRINT ALL STORED DATA ON ANY OF  601  DIF-
FERENT MINERALS.  AS AN AID IN MINERAL IDENTIFICATION
IT WILL SORT THROUGH ALL THE MINERALS STORED AND LIST
ONLY THOSE THAT FIT THE DATA SUPPLIED BY THE USER.  THE
PROGRAM WILL REQUEST YOUR DATA, SORT, AND PRINT OUT THE
FIRST 25 MINERALS HAVING SIMILAR PROPERTIES.

     A LIST OF PROPERTIES WITH PROPERTY NUMBERS AND
ACCEPTABLE DATA FOR THE SORT ROUTINE FOLLOWS.

#1 HARDNESS: MOHS' , 0 TO 10 SCALE.
#2 SPECIFIC GRAVITY: ANY NUMBER>0 AND <20
#3 CRYSTAL SYSTEM: TYPE THE FIRST THREE LETTERS ONLY OF ANY ONE
   OF THE SIX CRYSTAL SYSTEMS--ISO, HEX, ORT, TET, TRI, MON,
   OR CRY IF MINERAL IS CRYPTOCRYSTALLINE OR FIBROUS.
#4 OPTIC TYPE: GIVE THE FIRST THREE LETTERS OF ANY ONE OF THE
   THREE OPTIC TYPES--ISO, UNI, BIA, OR OPA FOR OPAQUE.
#5 REFRACTIVE INDEX: WHEN ASKED THE NUMBER OF INDICES TYPE
   1 IF THE MINERAL IS ISOTROPIC, 2 IF UNIAXIAL, 3 IF
   BIAXIAL.  THEN, WHEN ASKED, SUBMIT INDICES SEPARATED
   BY COMMAS.  IF ANY INDEX IS UNKNOWN, TYPE A ZERO IN
   ITS PLACE.
#6 OPTIC SIGN:  +OR - IS ACCEPTABLE

WOULD YOU LIKE TO (A) REFERENCE A PARTICULAR MINERAL,(B) SORT
FOR A MINERAL BY PROPERTIES, OR (S) STOP PROGRAM? B

TYPE PROPERTY NUMBERS? 1,2,3
TYPE HARDNESS? 6.5
TYPE SPECIFIC GRAVITY? 2.5
TYPE CRYSTAL SYSTEM? HEX

EUCRYPTITE,    KALIOPHILITE  KALSILITE     LEIFITE        MICROSOMMITE
NEPHELITE      QUARTZ

DO YOU HAVE MORE PROPERTIES FOR THE SAME MINERAL? YES

TYPE PROPERTY NUMBERS? 4
TYPE OPTIC TYPE? UNI

EUCRYPTITE     KALIOPHILITE  KALSILITE     LEIFITE        MICROSOMMITE
NEPHELITE      QUARTZ

DO YOU HAVE MORE PROPERTIES FOR THE SAME MINERAL? YES

TYPE PROPERTY NUMBERS? 5
HOW MANY INDICES? 2
TYPE IN EPSILON AND OMEGA? 1.556,0

QUARTZ

WOULD YOU LIKE TO (A) REFERENCE A PARTICULAR MINERAL,(B) SORT
FOR A MINERAL BY PROPERTIES, OR (S) STOP PROGRAM? A

TYPE MINERAL NAME? QUARTZ

QUARTZ          SPECIFIC GRAVITY:  2.6        HARDNESS:  7
CRYSTAL SYSTEM: HEX           OPTICAL TYPE: UNI              SIGN: +
INDEX:   1.553 <EPSILON< 1.553                 1.544 <OMEGA< 1.544
END POINTS INCLUDED

WOULD YOU LIKE TO (A) REFERENCE A PARTICULAR MINERAL,(B) SORT
FOR A MINERAL BY PROPERTIES, OR (S) STOP PROGRAM? S

TIME: 4.935 SEC.
READY
```

COST ESTIMATE

For the job included as the sample output, the total central-processor time was 4.9 seconds, at the current rate for Dartmouth of $0.11/sec. Total terminal time was about 7 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead
= $0.54 + $0.41 + communications + network overhead
= $0.95 + communications + network overhead

CONTENTS—MINERAL

pages
1 Identification & Abstract
3 User Instructions
5 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Mathematical Programming System
CALLING NAME	MPS/360
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary: available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

Linear programming (LP) is a mathematical technique designed to analyze the potentialities of alternate business activities and to choose those that permit the best use of resources in the pursuit of a desirable objective. It has many uses. For example, it can analyze capital, raw materials, manpower, plant and storage facilities and then translate its findings into minimum costs and maximum profits for its user; it can be used to allocate, assign, schedule, select or evaluate whatever possibilities limited resources possess for different jobs; it can blend, mix, distribute, control, order, budget, bid, cut, trim, price, purchase, and plan; it can deduce the most profitable method of transporting goods from plant to warehouse to outlet.

MPS/360 is composed of a set of procedures, a subset of which deals only with LP. The strategy for solving an LP problem is the ordered execution of a series of these procedures. A second set of procedures handles separable programming problems.

The user conveys the proposed strategy to MPS via the MPS control language. The procedure call statement of the control

continued

language calls the LP procedures and transfers arguments to them. MPS control statements are preprocessed by the control program COMPILER.

A catalogued procedure is available for simple linear programming models, reducing the user's work to merely data arrangement.

REFERENCES

Mathematical Programming System/360, Application Description (IBM Manual GH20-0136-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Control Language User's Manual (IBM Manual GH20-0290-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Linear and Separable Programming User's Manual (IBM Manual GH20-0476-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Message Manual (IBM Manual GH20-0603-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Copies of these manuals can be obtained through IBM branch offices.

USER INSTRUCTIONS

The catalogued procedure MPSGO is used to execute MPS/360 control programs. Moreover, it can be used without requiring a control program as input, if just a straightforward optimization is desired.

General Deck Setup

The Job Control Language required to compile and execute user supplied MPS decks is as follows:

```
JOB Card
//    EXEC   MPSGO
//MPS.SYSIN   DD     *

      (MPS   Control Program)

/*
//GO.SYSIN    DD     *

      (MPS   data deck)

/*
```

Deck Setup, Omitting Control Program

```
JOB Card
//    EXEC   MPSGO,MODE = MAX       )
                          MIN       }
                          MAXBOUND  }
                          MINBOUND  )
//GO.SYSIN    DD     *

      (MPS   data deck)

/*
```

On the above EXEC card, MODE must be assigned one of the following parameters:

MAX   , for LP maximization and no bounds on variables
MIN   , for LP minimization and no bounds on variables
MAXBOUND, for LP maximization with bounds
MINBOUND, for LP minimization with bounds

Vector names to be specified in data:

RHS   – right hand side vector name
BOUND – bound vector name (when used)
LINEAR – data set name on NAME card

continued

Restrictions

MPS/360 will handle problems having no more than 4094 constraints. A minimum partition size of 98K is required for execution of moderately-sized problems. Large problems will require a larger partition.

REFERENCES

Mathematical Programming System/360, Application Description (IBM Manual GH20-0136-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2, Control Language User's Manual (IBM Manual GH20-0290-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Version 2 Linear and Separable Programming User's Manual (IBM Manual GH20-0476-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Mathematical Programming System/360 (360A-CO-14X), Message Manual (IBM Manual GH20-0603-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

Copies of these manuals can be obtained through IBM branch offices.

SAMPLE INPUT

```
NAME            SAMPLE INPUT
ROWS
 N  VALUE
 N  VALUE1
 N  CHROW
 N  CHROW2
 N  CHROW3
 N  CHROW4
 E  YIELD
 L  FE
 L  CU
 L  MN
 L  MG
 G  AL
 L  SI
 N  CHROW5
COLUMNS
    BIN.1       VALUE           .03000      CHROW     -     .02000
    BIN.1       YIELD          1.00000      FE              .15000
    BIN.1       CU              .03000      MN              .02000
    BIN.1       MG              .02000      AL              .70000
    BIN.1       SI              .02000      VALUE1    -     .03000
    BIN.2       VALUE           .08000      CHROW           .01000
    BIN.2       CHROW2          .01000      YIELD          1.00000
    BIN.2       FE              .04000      CU              .05000
    BIN.2       MN              .04000      MG              .03000
    BIN.2       AL              .75000      SI              .06000
    BIN.2       VALUE1    -     .08000
    BIN.3       VALUE           .17000      CHROW4          .00500
    BIN.3       YIELD          1.00000      FE              .02000
    BIN.3       CU              .08000      MN              .01000
    BIN.3       AL              .80000      SI              .08000
    BIN.3       VALUE1    -     .17000
    BIN.4       VALUE           .12000      CHROW4    -     .01000
    BIN.4       YIELD          1.00000      FE              .04000
    BIN.4       CU              .02000      AL              .75000
    BIN.4       MN              .02000      SI              .12000
    BIN.4       VALUE1    -     .12000
    BIN.5       CHROW2    -     .00500      CHROW3    -     .01000
    BIN.5       YIELD          1.00000      FE              .02000
    BIN.5       VALUE           .15000      CU              .06000
    BIN.5       MN              .02000      MG              .01000
    BIN.5       AL              .80000      ST              .02000
    BIN.5       VALUE1    -     .15000
```

continued

```
    ALUM      VALUE           .21000   CHROW3          .01000
    ALUM      YIELD          1.00000   FE              .01000
    ALUM      CU              .01000   AL              .97000
    ALUM      SI              .01000   VALUE1    -     .21000
    SILCON    VALUE           .38000   YIELD          1.00000
    SILCON    FE              .03000   SI              .97000
    SILCON    VALUE1    -     .38000
RHS
    ZZZZZZ04  YIELD       2000.00000   FE            60.00000
    ZZZZZZ04  CU           100.00000   MG            30.00000
    ZZZZZZ04  MN            40.00000   SI           300.00000
    ZZZZZZ04  AL          1500.00000
    ZZZZZZ05  VALUE     -     .01000   FE              .01000
    ZZZZZZ05  AL        -     .01000
    ZZZZZZ06  CU        -    1.00000   MN             2.00000
    ZZZZZZ07  FE             3.00000   MN        -    1.00000
    ZZZZZZ08  AL            25.00000   SI        -    5.00000
    ZZZZZZ09
RANGES
    ZZZZZZ04  SI            50.00000
BOUNDS
 UP ZZZZZZ01  BIN.1        200.00000
 UP ZZZZZZ01  BIN.2       2500.00000
 LO ZZZZZZ01  BIN.3        400.00000
 UP ZZZZZZ01  BIN.3        800.00000
 LO ZZZZZZ01  BIN.4        100.00000
 UP ZZZZZZ01  BIN.4        700.00000
 UP ZZZZZZ01  BIN.5       1500.00000
ENDATA
```

SAMPLE OUTPUT

```
SETUP  PBFILE

TIME =    0.16

 MATRIX1  ASSIGNED TO MATRIX1

 ETA1     ASSIGNED TO ETA1

 SCRATCH1 ASSIGNED TO SCRATCH1
 SCRATCH2 ASSIGNED TO SCRATCH2

 MAXIMUM PRICING NOT REQUIRED - MAXIMUM POSSIBLE   7

 NO CYCLING

      POOLS          NUMBER   SIZE    CORE
 H.REG-BIRS MAP                         68
 WORK   REGIONS        9       136    1224
 MATRIX BUFFERS        2       928    1856
 ETA    BUFFERS        4      2288    9152

                           TOTAL  NORMAL  .FREE.  FIXED BOUNDED
 ROWS    (LOG.VAR.)          14      6       7      1      0
 COLUMNS (STR.VAR.)           7      7       0      0      0

       77 ELEMENTS - DENSITY =  26.19 -   2 MATRIX RECORDS (WITHOUT RHS'S)

PRIMAL        OBJ = VALUE        RHS = ZZZZZZ04

TIME =     0.49  MINS.        PRICING    7
 SCALE =     .

   ITER   NUMBER  VECTOR  VECTOR  REDUCED       SUM
  NUMBER  INFEAS    OUT     IN      COST       INFEAS
M    1      0        7      20    1.97000-       .
 FEASIBLE SOLUTION

PRIMAL        OBJ = VALUE        RHS = ZZZZZZ04

TIME =     0.51  MINS.        PRICING    7
 SCALE =     .
 SCALE RESET TO    1.00000

   ITER   NUMBER  VECTOR  VECTOR  REDUCED   FUNCTION
  NUMBER  NONOPT    OUT     IN      COST      VALUE
M    2      5       10      16    .13000-  290.000
     3               8      18    .02500-  273.333
     4              20      19    .00333-  270.000
 OPTIMAL SOLUTION
```

COST ESTIMATE

The above sample problem having 7 constraints and 7 original variables with upper bound restrictions on the variables is discussed in User Instructions Ref. 3. This sample problem was executed in class A (100K high speed core) and required 0.57 minutes of central processor unit time, at a cost of $4.80.

Charge to user = S/360 charges + handling/consultation charges + postage + network overhead

= $4.80 + handling/consultation charges + postage + network overhead

CONTENTS—MPS/360

pages	
1– 2	Identification & Abstract
3– 4	User Instructions
5– 7	I/O
9	Cost—Contents

DESCRIPTIVE TITLE	Univariate and Multivariate Analysis of Variance, Covariance and Regression
CALLING NAME	NYBMUL
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Jeremey D. Finn State University of New York at Buffalo
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

NYBMUL performs an exact least-squares multivariate analysis of variance or covariance for any crossed and/or nested design. The number of observations per cell may be equal, proportional or disproportionate, including missing observations and incomplete designs. The program can also be used to perform regression analysis, canonical correlation and discriminant analysis.

Input may be either the raw data or a variance-covariance matrix together with means and frequencies. A variety of data transformations are provided which include the use of a matrix transformation. Estimation and analysis phases are based on contrasts which may be specified by the user.

REFERENCES

NYBMUL: Univariate and Multivariate Analysis of Variance and Covariance (Buffalo, N.Y.: SUNY at Buffalo, Comput. Center, 1968).

NYBMUL Manual (Buffalo, N.Y.: Comp. Center Press, SUNY at Buffalo, 1969).

USER INSTRUCTIONS

The user must supply a Program Control Card Set to specify the analysis to be performed. If the data are on cards, they will be included in this set. A NYBMUL instruction manual[1], available from the State University of New York at Buffalo, is required for the construction and arrangement of this input.

General Deck Setup

Cards 1—5 replace the INITIAL SYSTEM CONTROL CARDS listed in paragraph A on page 7 and described on pages 8—9 of Ref. 1.

Card 1

Standard JOB Card (Minimum partition=142K)

Card 2

```
// EXEC NYBMUL
```

Card 3 (needed if input data are on tape or disk rather than cards)

```
//GO.FTxxF001 DD----
```

where xx is replaced by the data set reference number of the input data (same number as punched in Cols. 35-36 of the INPUT DESCRIPTION CARD). This number must not be 0, 1,2,3,4,5,6,7, or 10, and must be less than 30.
---- is replaced by the appropriate information

Card 4 (needed for designs with a large number of non-empty cells in order to obtain additional disk work space)

```
//GO.FT03F001  DD  SPACE=(TRK,(c,10)),DCB=(LRECL=a,BLKSIZE=b)
```

The user can determine if this card is needed by computing:

n = 4 * NFULL, where NFULL is the number of non-empty cells

If n > 500, this card should be included, supplying the following integer specifications.

a = n + 4
b = n + 8
c = (NRANK) ÷ x , rounded to nearest integer

continued

where NRANK is the rank of the model for significance testing and x is the number of blocks per track as can be estimated from the following table:

b	880	1600	3072	3520	3625	7294
x	7	4	2	2	1	1

Experience indicates that such problems, where $n > 500$, tend to require large amounts of CPU time. In addition, when $n > 600$ the minimum partition size may increase beyond 150K.

Card 5

```
//GO.SYSIN  DD  *
```

Program Control Card Set

These cards are described in the NYBMUL manual[1], starting at page 10. There are two deviations from that description.

1. The version at Washington University reads only the first four characters of the six character labels on the Factor Identification Cards (page 13), and on the Variable Label Cards (page 18). Thus it is recommended that the user supply labels having a maximum of four characters, left-justified in the appropriate fields.

2. The two Final System Control Cards shown in paragraph C on pages 7 and 41 are to be omitted from this card set.

Card 6

```
/*
```

REFERENCES

1. *NYBMUL Manual* (Buffalo, N.Y.: Comp. Center Press, SUNY at Buffalo, 1969).

2. *NYBMUL: Univariate and Multivariate Analysis of Variance and Covariance* (Buffalo, N.Y.: SUNY at Buffalo, Computer Center, 1968).

SAMPLE INPUT

```
//   STANDARD JOB CARD (SEE OS NOTEBOOK)(MINIMUM PARTITION=142K)
// EXEC NYBMUL
//GO.SYSIN DD *
2 X 2 X 2 FACTORIAL DESIGN,   EQUAL N'S
ALLEN L. EDWARDS,'EXPERIMENTAL DESIGN', PAGE 219
   1   3   1   1
A         2B         2C        2
FINISH
(3(I1,2X),F2.0)
X
1  1  1  76
1  1  1  66
1  1  1  43
1  1  1  62
1  1  1  65
1  1  1  43
1  1  1  42
1  1  1  60
1  1  1  78
1  1  1  66
1  1  2  36
1  1  2  45
1  1  2  47
1  1  2  23
1  1  2  43
1  1  2  43
1  1  2  54
1  1  2  45
1  1  2  41
1  1  2  40
1  2  1  43
1  2  1  75
1  2  1  66
1  2  1  46
1  2  1  56
1  2  1  62
1  2  1  51
1  2  1  63
1  2  1  52
1  2  1  50
1  2  2  37
1  2  2  22
1  2  2  25

   :
   :

2  1  1  80
```

continued

```
2  1  1  63
2  1  1  58
2  1  2  74
2  1  2  74
2  1  2  64
2  1  2  86
2  1  2  68
2  1  2  72
2  1  2  62
2  1  2  64
2  1  2  78
2  1  2  61
2  1  2  67
2  2  1  64
2  2  1  70
2  2  1  65
2  2  1  60
2  2  1  55
2  2  1  57
2  2  1  66
2  2  1  79
2  2  1  80
2  2  2  67
2  2  2  60
2  2  2  54
2  2  2  51
2  2  2  49
2  2  2  38
2  2  2  55
2  2  2  56
2  2  2  68
2  2  2  58

    8                                                  1
1,2,3,1*2,1*3,2*3.
C0,C0,C0,    GM
C1,C0,C0     A
C0,C1,C0     B
C0,C0,C1     C
C1,C1,C0     AB
C1,C0,C1     AC
C0,C1,C1     BC
C1,C1,C1     ABC
    1
-1,1,1,1,1,1,1,1.

/*                                                                        STOP
```

SAMPLE OUTPUT

```
                  * * * * * * * * * M U L T I V A R I A N C E * * * * * * * * *
                   *                                                          *
                    *      UNIVARIATE AND MULTIVARIATE ANALYSIS OF          *
                     *      VARIANCE, COVARIANCE AND REGRESSION           *
                      *      VERSION 4           JUNE 1968              *
                       *                                              *
                        * * * * * * * * * * * * * * * * * * * *


                                     PROBLEM
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
2 X 2 X 2 FACTORIAL DESIGN,   ECUAL N'S              ALLEN L. EDWARDS,'EXPERIEMENTAL DESIGN', PAGE 219
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

                                     COMMENTS



                                 INPUT PARAMETERS
                                 ================


                  NUMBER OF VARIABLES IN INPUT VECTORS=           1

                  NUMBER OF FACTORS IN DESIGN=                    3

                    NUMBER OF LEVELS OF FACTOR 1 (A   ) =      2
                    NUMBER OF LEVELS OF FACTOR 2 (B   ) =      2
                    NUMBER OF LEVELS OF FACTOR 3 (C   ) =      2

               INPUT FROM CARD. OPTION 1

                  FORMAT OF DATA
                     (3(I1,2X),F2.0)
  1ST CBS
  76.00C0
```

continued

```
                                  CELL IDENTIFICATION AND FREQUENCIES

CELL FACTOR-LEVELS        N
          A     B     C
   1      1     1     1   10
   2      1     1     2   10
   3      1     2     1   10
   4      1     2     2   10
   5      2     1     1   10
   6      2     1     2   10
   7      2     2     1   10
   8      2     2     2   10

                                        TOTAL N=     80.

2 X 2 X 2 FACTORIAL DESIGN,   EQUAL N'S              ALLEN L. EDWARDS,'EXPERIEMENTAL DESIGN', PAGE 219

                         OBSERVED CELL MEANS --- ROWS ARE CELLS-COLUMNS ARE VARIABLES
                         ------------------------------------------------------------

              1
              X

   1      60.09999
   2      41.70000
   3      56.39999
   4      24.70000
   5      77.00000
   6      70.29999
   7      66.29999
   8      55.59999

                         OBSERVEC CELL STD DEVS--ROWS ARE CELLS-COLLMNS VARIABLES
                         ------------------------------------------------------------

              1
              X

   1      13.26195
   2       8.09736
   3       9.94658
   4       6.68416
   5      10.67708
   6       8.00098
   7       8.32750
  .8       8.73317
```

continued

```
                                   OBSERVED  COMBINED MEANS
                                   ===========================
-------------------------------------------------------------------------------------
FACTORS    1 (A   )                 FACTORS    2 (B   )                 FACTORS    3 (C   )

  LEVEL         1                     LEVEL         1                     LEVEL         1
  N =    40.                          N =    40.                          N =    40.

    MEANS          X   =  45.72         MEANS          X   =  62.27         MEANS          X   =  64.95
    -------                             -------                             -------

  LEVEL         2                     LEVEL         2                     LEVEL         2
  N =    40.                          N =    40.                          N =    40.

    MEANS          X   =  67.30         MEANS          X   =  50.75         MEANS          X   =  48.07
    -------                             -------                             -------
-------------------------------------------------------------------------------------
FACTORS    1 (A   )  2 (B   )         FACTORS    1 (A   )  3 (C   )         FACTORS    2 (B   )  3 (C   )

  LEVEL         1          1          LEVEL         1          1          LEVEL         1          1
  N =    20.                          N =    20.                          N =    20.

    MEANS          X   =  50.90         MEANS          X   =  58.25         MEANS          X   =  68.55
    -------                             -------                             -------

  LEVEL         1          2          LEVEL         1          2          LEVEL         1          2
  N =    20.                          N =    20.                          N =    20.

    MEANS          X   =  40.55         MEANS          X   =  33.20         MEANS          X   =  56.00
    -------                             -------                             -------

  LEVEL         2          1          LEVEL         2          1          LEVEL         2          1
  N =    20.                          N =    20.                          N =    20.

    MEANS          X   =  73.65         MEANS          X   =  71.65         MEANS          X   =  61.35
    -------                             -------                             -------

  LEVEL         2          2          LEVEL         2          2          LEVEL         2          2
  N =    20.                          N =    20.                          N =    20.

    MEANS          X   =  60.95         MEANS          X   =  62.95         MEANS          X   =  40.15
    -------                             -------                             -------
-------------------------------------------------------------------------------------
```

continued

```
                              ESTIMATION PARAMETERS
                              =====================


          RANK OF THE BASIS = RANK OF MODEL FOR SIGNIFICANCE TESTING =  8
          RANK OF THE MODEL TO BE ESTIMATED IS  0
          ERROR TERM TO BE USED IS (WITHIN CELLS)










                              SYMBOLIC CONTRAST VECTORS
                              =========================

CO,CO,CO,                                                           GM
C1,CO,CO,                                                           A
CO,C1,CO,                                                           B
CO,CO,C1,                                                           C
C1,C1,CO,                                                           AB
C1,CO,C1,                                                           AC
CO,C1,C1,                                                           BC
C1,C1,C1,                                                           ABC
```

continued

```
2 X 2 X 2 FACTORIAL DESIGN,   EQUAL N'S                 ALLEN L. EDWARDS,'EXPERIEMENTAL DESIGN', PAGE 219

                                        SAMPLE CORRELATION MATRIX
                        ------------------------------------------------------------

                  1
                  X

1  X        1.000000




                                VARIABLE        VARIANCE        STANDARD DEVIATION

                                1 X             88.587814              9.4121

                                                D.F.=    72.
                                ERROR TERM FOR ANALYSIS OF VARIANCE WITHIN CELLS



                                                ANALYSIS OF VARIANCE
                                                ====================
                                                ====================

                                                1 DEPENDENT VARIABLES
                                                       1 X
```

continued

```
                                   HYPOTHESIS   1         1 DEGREES OF FREEDOM
                                   ===========================================

             CO,CO,CO,                                                              GM
                      - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

                                      TESTS OF HYPOTHESIS BEING SKIPPED
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
                                   HYPOTHESIS   2         1 DEGREES OF FREEDOM
                                   ===========================================

             C1,CO,CO,                                                              A
                      - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

                                   UNIVARIATE ANALYSIS OF VARIANCE FOR (X   )
  HYPOTHESIS MEAN SQUARE=     9309.6016     F=     105.0889     WITH     1. AND      72. DEGREES OF FREEDOM      P LESS THAN 0.0001
                                         ERROR TERM IS ON PAGE    5
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
                                   HYPOTHESIS   3         1 DEGREES OF FREEDOM
                                   ===========================================

             CO,C1,CO,                                                              B
                      - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

                                   UNIVARIATE ANALYSIS OF VARIANCE FOR (X   )
  HYPOTHESIS MEAN SQUARE=     2656.5093     F=      29.9873     WITH     1. AND      72. DEGREES OF FREEDOM      P LESS THAN 0.0001
                                         ERROR TERM IS ON PAGE    5
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
                                   HYPOTHESIS   4         1 DEGREES OF FREEDOM
                                   ===========================================

             CO,CO,C1,                                                              C
                      - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

                                   UNIVARIATE ANALYSIS OF VARIANCE FOR (X   )
  HYPOTHESIS MEAN SQUARE=     5695.3047     F=      64.2899     WITH     1. AND      72. DEGREES OF FREEDOM      P LESS THAN 0.0001
                                         ERROR TERM IS ON PAGE    5
```

continued

```
000 0094
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
                                        HYPOTHESIS   5          1 DEGREES OF FREEDOM
                                        ==========================================

              C1,C1,C0,                                                                     AB
                 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
                                        UNIVARIATE ANALYSIS OF VARIANCE FOR (X   )
   HYPOTHESIS MEAN SQUARE=       27.6130      F=      0.3117      WITH     1. AND      72. DEGREES OF FREEDOM      P LESS THAN 0.5784
                                                   ERROR TERM IS ON PAGE    5
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
                                        HYPOTHESIS   6          1 DEGREES OF FREEDOM
                                        ==========================================

              C1,C0,C1,                                                                     AC
                 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
                                        UNIVARIATE ANALYSIS OF VARIANCE FOR (X   )
   HYPOTHESIS MEAN SQUARE=     1336.6064      F=     15.0879      WITH     1. AND      72. DEGREES OF FREEDOM      P LESS THAN 0.0003
                                                   ERROR TERM IS ON PAGE    5
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
                                        HYPOTHESIS   7          1 DEGREES OF FREEDOM
                                        ==========================================

              C0,C1,C1,                                                                     BC
                 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
                                        UNIVARIATE ANALYSIS OF VARIANCE FOR (X   )
   HYPOTHESIS MEAN SQUARE=      374.1096      F=      4.2230      WITH     1. AND      72. DEGREES OF FREEDOM      P LESS THAN 0.0436
                                                   ERROR TERM IS ON PAGE    5
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
                                        HYPOTHESIS   8          1 DEGREES OF FREEDCM
                                        ==========================================

              C1,C1,C1,                                                                     ABC
                 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
                                        UNIVARIATE ANALYSIS OF VARIANCE FOR (X   )
   HYPOTHESIS MEAN SQUARE=      108.1125      F=      1.2204      WITH     1. AND      72. DEGREES OF FREEDOM      P LESS THAN 0.2730
                                                   ERROR TERM IS ON PAGE    5
* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
000 0094
```

COST ESTIMATE

The sample problem, executed in class F (150K low speed core), required 0.5 minutes of central processor unit time, at a cost of $1.60.

Charge to user = S/360 charges + handling/consultation charges + postage + network overhead
= $1.60 + handling/consultation charges + postage + network overhead

CONTENTS—NYBMUL

pages

1	Identification & Abstract
3– 4	User Instructions
5–13	I/O
15	Cost—Contents

DESCRIPTIVE TITLE	University of Maryland Test Scoring Program (Version 4)
CALLING NAME	UOM 32
INSTALLATION NAME	University of Maryland Computer Science Center
AUTHOR(S) AND AFFILIATION(S)	Dr. George Green University of Maryland
LANGUAGE	FORTRAN IV
COMPUTER	UNIVAC 1108
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Sam Steinberg, Program Librarian Computer Science Center, University of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This system scores objective tests recorded on Digitek Optical Reader Sheets. Using an answer key it scores tests of up to 160 items for any number of students.

Output

1. A listing of students' answers to each question is optional.
2. For each student, the program lists the number of questions answered right, the number wrong, and the number omitted. Optionally included is a score which imposes a penalty for wild guessing. A standardized T score (which has an arithmetic mean of 50 and a standard deviation of 10) is also printed for each student.
3. A frequency distribution (histogram) and the mean and standard deviation of the scores are printed.
4. An item analysis is printed for *each question* on the examination.
    a) P (Difficulty Index) is the proportion of the total group who answer the question correctly. Items having a

continued

difficulty index of .50 provide maximum differentiation among students. Good tests contain items having difficulty indices in the middle range (.25 to .75).

b) D (Discrimination Index) measures the power of a single item to discriminate between the upper and lower halves of the student group, divided as to total scores. A good classroom test should have indices of discrimination of .30 or better.

Also printed are the responses to each item by the high and low scoring groups. The correct answer is indicated.

The three statistics which follow the D index are used for similar purposes and are consistent in meaning with it. All are discrimination indices expressing the relationship between item success and the total score (the criterion).

c) PHI (Phi coefficient) is an index of discriminating power when the criterion variable is a natural dichotomy and must be used as such; i.e., high-low, good-poor, etc. No assumptions are made about the form of the distribution of the group.

d) RPBI (Point biserial correlation coefficient) is an index of discriminating power when the criterion is a continuous variable. No assumptions are made about the form of the distribution of the group.

e) RB (Biserial correlation coefficient) is an index of discriminating power requiring the assumption that one of the normally distributed underlying variables has been forced into a dichotomy.

The choice among these depends partly on the purpose for which the test and item analysis data are to be used, and partly on the convenience with which each statistic serves that purpose. For most classroom examinations, the D statistic is most easily understood; it will identify items with little internal-consistency discriminating power and lead to greater efficiency of measurement when revising a test to contain the more discriminating items. It should be noted, however, that when RB and RPBI are used, the indices are usually not equal. RB tends to be substantially larger than RPBI.

continued

f) An overall measure of test reliability is provided by Kuder Richardson No. 20 and No. 21 statistics. This is an estimate of how close the same set of scores would result if the same set of items were given again. Most test writers settle with reliabilities over .60 for teacher-made tests. An index of .80 and above reveals a highly reliable examination.

Formulae

Phi Coefficient

$$\emptyset = \frac{BC - AD}{[(A + B)\ (C + D)\ (A + C)\ (B + D)]^{0.5}}$$

Point Biserial Correlation

$$r_{pbi} = \frac{\overline{X}p - \overline{X}_t}{S_t} \sqrt{\frac{p}{q}}$$

$\overline{X}$ = the mean of all scores

$\overline{X}p$ = the mean of all scores in the upper half

$\overline{X}_t$ = the mean of all scores in the lower half

p = proportion of individuals in upper half

q = proportion of individuals in lower half

S_t = standard deviation of all scores

Biserial Correlation

$$r_{bi} = \frac{\overline{X}_p - \overline{X}_t}{S_t} \times \frac{P}{y}$$

P = proportion of cases in the upper group X proportion of cases in the lower group; $P = pq$

y = height of ordinate of unit normal curve at point of division between p and q

Discrimination Index

$$DIS = \frac{X - Z}{N} \times 2$$

continued

X = total correct answers in upper half

Z = total correct answers in lower half

N = sample size

Kuder-Richardson 20

$$\text{K–R (20)} = \frac{NQ}{NQ-1}\,\frac{(S_t^2 - PQ)}{S_t^2}$$

Kuder Richardson 21

$$\text{K–R (21)} = \frac{NQ}{NQ-1}\left[1 - \frac{\overline{X}_t(NQ-\overline{X}_t)}{NQ\ S_t^2}\right]$$

NQ = the number of questions

PQ = p times q summed over all questions

Correction for Guessing

$$\text{Raw Score} = R - \frac{W}{A-1}$$

R = Rights

W = Wrongs

A = Number of options per item (2 to 5)

REFERENCES

Ebel, R.L., *Measuring Educational Achievements*, (Prentice-Hall, Inc., New Jersey, 1965).

Lindquist, E.F., (Ed.), *Educational Measurement*, (American Council on Education, Washington D.C., 1951).

USER INSTRUCTIONS

Tests must be administered on Digitek Optical Reader Test Scoring Sheets. These may be obtained from the University of Maryland Student Union Book Store or directly from the Optical Scanning Corporation, Post Office Box 40, Newtown, Pennsylvania.

The following conditions are important in the administration of the test:

a) Students should use ordinary pencil, number 2 preferred. Erasures should be completely clean. *DO NOT* use pens or electrographic pencils.

b) Each student should have a unique number (upper right corner of sheet), as the output from the scoring program uses that number for identification purposes. Many people use the number in the grade book or ask the students to use their University identification number. The identification section of the sheet must *not* be blank, or it will be skipped.

c) It may be helpful to add the Section number after the student identification number (space 1). The forms have a total of ten locations for coding in test identification numbers.

Student Response Cards

The Computer Science Center will use the Test Scoring Sheets to generate the deck of Student Response Cards.

Columns	*Contents*
1–10	Student Number
12	Card Sequence Letter (J for first card, K, L, and M for second, third, and fourth cards)
13–52	Student Responses to test items 1–40 on Card J ... 121–160 on Card M.

Note: A blank for a particular question on the answer key causes this question to be disregarded in the computation of student scores.

The Answer Key Cards are normally generated as output from the Digitek Optical Scanner. Simply put the correct answers on an answer sheet with zeros in all the identification columns. This sheet must be submitted along with student answer sheets, to keypunch for Digitek processing.

continued

Instruction Card

Columns	*Contents*
4	Number of possible answers to each question $\leq$ 5)
6– 8	Number of questions in the test ($\leq$ 160)
12	0: Do not list printed output by sections. 1: List printed output by sections.
16	0: No penalty for wild guessing. 1: Penalty for wild guessing.
20	0: No exams have questions in an alternative order. 1: Some exams have questions in an alternative order.
24	Alternate test code number. (Fill in if Col. 20 is punched 1. This same number must be punched in Col. 80 of all alternate exams.)
28	Test scores for each student should be: 0: printed only 1: printed and punched on cards 2: printed and written on logical tape drive unit B3
32	0: No printed output of student responses is desired. 1: Printed output of student responses is desired.
36	0: Read a Title Card and answer key. 1: Use the same Title Card and answer key as the previous job.

Title Card

The Title Card must be provided if Col. 36 of the Instruction Card is punched 0. It contains up to 80 columns of any alphanumeric information desired by the user.

Alternative Order Cards

The Alternative Order Cards must be provided if Col. 36 of the Instruction Card is punched 1. Up to eight cards may be used.

Columns	*Contents*
1– 4 (first card)	Number of item on original test which is first item on alternate test.
5– 8 (first card)	Number of item on original test which is second item on alternate test.
⋮	⋮

continued

Columns	*Contents*
77–80 (eighth card)	Number of item on original test which is 160th item on alternate test.

Note: If an alternative examination is used, all answer sheets should first be sorted into main exam sheet and alternative exam sheets before Digitek Processing.

Answer Key Cards

Answer Key Cards must be provided if Col. 36 of the Instruction Card is punched 0. They may be punched or generated from a Test Scoring Sheet with all zeros as the Student Number.

Columns	*Contents*
1–11	00000000000
12	Card Number (J, K, L, or M)
13–52	Correct answers to items 1–40 on Card J ... 121–160 on Card M.

End of Job Card

Columns	*Contents*
1–10	0000000000

If more than one job is to be run at the same time, simply follow the End of Job Card for the first run with the Instruction and Data Cards for the second run, etc.

SAMPLE INPUT

```
'RUN  UOM32,001-02-007,STEINBERG
'XOT RMDUOM*LIBRARY.UOM32
   5 160                        1
1   THIS IS A TEST OF  UOM32
0000000000 J4200100103144331241030312433123213120122
0000000000 K0212122320022021311013032310133211030010
0000000000 L24230423 0412333232000203113210310231233
0000000000 M   310103201212310
5776095281 J4200100103144331241030312433123213120122
5776095281 K0212122320022021311013032310133211030010
5776095281 L24230423 0412333232000203113210310231233
5776095281 M 013101032012123101113231031
1192648911 J4303100103444313040022222432221013120322
1192648911 K32121333302222223121330423100 3201031223
1192648911 L33232323 3432333331310213313320030231333
1192648911 M 033131222002223100120321330
1623294691 J1200100133304333211120212401213213120122
1623294691 K3310132031320221300313022330133210032211
1623294691 L34232423 0110233032003201103022303233333
1623294691 M 013332022303123101102211330
2174070361 J4233120013401203413124242400213013120122
2174070361 K3212132031321221302313032313043213023211
2174070361 L03230320 3410131012303233113000321231203
2174070361 M 013301222323223102103211210
5783074891 J4431120133444432243220212331033213320322
5783074891 K32101223303222013023130023113 3210030233
5783074891 L24332111 0330323032003301113020120331233
5783074891 M 013101021022223301100231330
1973051301 J4400  01343443143421 2113431002210 20112
1973051301 K1210132231222222 03013222310 3102100220
1973051301 L33232423 2410333032020203013310310231133

     .                .
     .                .
     .                .

4096684071 J4403130103344231442101202432212013120022
4096684071 K12131023003223213231130223103 0232032303
```

continued

000 0095

000 0095

```
4096684071 L04232321 0412333012210003113212303030213
4096684071 M 133032122303023203100112332
4843494441 J4233100103344314341122112330012013120022
4843494441 K32101333202222223331130423113 3011032223
4843494441 L24232413 0110313022003203113020320211233
4843494441 M 0131012221122223201112231330
0943204581 J4430130103144313420102112303232013120122
0943204581 K3212122320222221322113042300030133032323
0943204581 L33232111 0112333032100213113001333212230
0943204581 M 033111322302323101100330332
2154251471 J1413140011324310243042232333313013030121
2154251471 K42022332222220223000131220123 3211033212
2154251471 L24232110 0110333233002103312123333211331
2154251471 M 030120222332123233122331230
2194212881 J4033122134144113211030202433212043323222
2194212881 K4002122330233001003213141032130131330013
2194212881 L02032412 3012333032202200313022020210132
2194212881 M 131131123303223233130021213
5793849191 J4403110103111130221122022401223213120122
5793849191 K2300122321223222322313331301130121322223
5793849191 L34232423 3110313012003200313320130211133
5793849191 M 013101312202223201110231330
5775022241 J4211120134344323140013311311301012010121
5775022241 K3232122330222221322313132310033203332023
5775022241 L32020410 0410310032011203112220123231233
5775022241 M 033031232122003211001001010
000000000
```

000 0095

SAMPLE OUTPUT

```
   THIS IS A TEST OF  UOM32


       ----  UNIVERSITY  OF  MARYLAND  ----

  TEST SCORING PROGRAM.... VERSION 2 .... 3/25/66

  (ORIGINAL SOURCE-TRAINING FOR BUSINESS AND INDUSTRY
  VOL. 2, NO.4, JULY-AUG. 1965. MODIFIED BY
  GEORGE R. GREEN, ECONOMICS DEPT.


NO. RESPONSES 5
NO. QUESTIONS 160
SECTION PRINTOUT 0
SUBTRACT CHANCE 0
ALTER. ORDER 0
ALTER. EXAM CODE 0
SUMMARY PUNCH 0
PRINT IND. QUEST. ANS. 1
DELETE READING OF TITLE AND KEY 0

KEY,            5 3 1 1 2 1 1 2 1 4 2 5 5 4 4 2 3 5 2 1 4 1 4 2 3 5 4 4 2 3 4 3 2 4 2 3 1 2 3 3
                1 3 2 3 2 3 3 4 3 1 1 3 3 1 3 2 4 2 2 1 2 4 1 4 3 4 2 1 2 4 4 3 2 2 1 4 1 1 2 1
                3 5 3 4 1 5 3 4 0 1 5 2 3 4 4 4 3 4 3 1 1 1 3 1 4 2 2 4 3 2 1 4 2 1 3 4 2 3 4 4
                0 0 0 4 2 1 2 1 4 3 1 2 3 2 3 4 2 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
```

continued

```
THIS IS A TEST OF  UON32

577609528   1     5 3 1 1 2 1 1 2 1 4 2 5 5 4 4 2 3 5 2 1 4 1 4 2 3 5 4 4 2 3 4 3 2 4 2 3 1 2 3 3 1 3 2 3 2 3 3 4 3 1
                  1 3 3 1 3 2 4 2 2 1 2 4 1 4 3 4 2 1 2 4 4 3 2 2 1 4 1 1 2 1 3 5 3 4 1 5 3 4 0 1 5 2 3 4 4 4 3 4 3 1
                  1 1 3 1 4 2 2 4 3 2 1 4 2 1 3 4 2 3 4 4 0 1 2 4 2 1 2 1 4 3 1 2 3 2 3 4 2 1 2 2 2 4 3 4 2 1 4 2 0 0
                  0 0 0 0 0 0 0 0 0 0

119264891   1     5 4 1 4 2 1 1 2 1 4 5 5 5 4 2 4 1 5 1 1 3 3 3 3 3 5 4 3 3 3 2 1 2 4 2 3 1 4 3 3 4 3 2 3 2 4 4 4 4 1
                  3 3 3 3 3 3 4 2 3 2 4 4 1 5 3 4 2 1 1 0 4 3 1 2 1 4 2 3 3 4 4 4 3 4 3 4 3 4 0 4 5 4 3 4 4 4 4 4 2 4
                  2 1 3 2 4 4 2 4 4 3 1 1 4 1 3 4 2 4 4 4 0 1 4 4 2 4 2 3 3 3 1 1 3 3 3 4 2 1 1 2 3 1 4 3 2 4 4 1 0 0
                  0 0 0 0 0 0 0 0 0 0

162329469   1     2 3 1 1 2 1 1 2 4 4 4 1 5 4 4 4 3 2 2 2 3 1 3 2 3 5 1 2 3 2 4 3 2 4 2 3 1 2 3 3 4 4 2 1 2 4 3 1 4 2
                  4 3 1 3 3 2 4 1 1 4 2 4 1 3 3 4 4 1 2 4 4 3 2 1 1 4 3 3 2 2 4 5 3 4 3 5 3 4 0 1 2 2 1 3 4 4 1 4 3 1
                  1 4 3 1 2 2 1 4 1 3 3 4 1 4 3 4 4 4 4 4 0 1 2 4 4 4 3 1 3 3 4 1 4 2 3 4 2 1 2 2 1 3 3 2 2 4 4 1 0 0
                  0 0 0 0 0 0 0 0 0 0

217407036   1     5 3 4 4 2 3 1 1 2 4 5 1 2 3 1 4 5 2 4 2 3 5 3 5 3 5 1 1 3 2 4 1 2 4 2 3 1 2 3 3 4 3 2 3 2 4 3 1 4 2
                  4 3 2 3 3 2 4 1 3 4 2 4 1 4 3 4 2 4 1 5 4 3 2 4 1 3 4 3 2 2 1 4 3 4 1 4 3 1 0 4 5 2 1 2 4 2 1 2 3 4
                  1 4 3 4 4 2 2 4 1 1 1 4 3 2 3 4 2 3 1 4 0 1 2 4 4 1 2 3 3 3 4 3 4 3 3 4 2 1 3 2 1 4 3 2 2 3 2 1 0 0
                  0 0 0 0 0 0 0 0 0 0

    .                                      .                                          .
    .                                      .                                          .
    .                                      .                                          .

215425147   1     2 5 2 4 2 5 1 1 2 2 4 3 5 4 2 1 3 5 4 1 5 3 3 4 3 4 4 4 4 2 4 1 2 4 1 4 1 2 3 2 5 3 1 3 3 4 4 3 3 3
                  3 3 3 1 3 3 4 1 1 1 2 4 2 3 3 1 2 3 4 0 4 3 2 2 1 4 4 3 2 3 3 5 3 4 3 2 2 1 0 1 2 2 1 4 4 4 3 4 4 1
                  1 3 2 1 4 4 2 3 2 3 4 4 4 4 3 2 2 4 4 2 0 1 4 1 2 3 1 3 3 3 4 4 3 2 3 4 3 4 4 2 3 3 4 4 2 3 4 1 0 0
                  0 0 0 0 0 0 0 0 0 0

219421288   1     5 1 4 4 2 3 3 2 4 5 2 5 5 2 2 4 3 2 2 1 4 1 3 1 3 5 4 4 3 2 3 1 5 4 4 3 4 3 3 3 5 1 1 3 2 3 3 4 4 1
                  3 4 4 1 1 2 1 1 4 3 2 4 2 5 2 1 4 3 2 4 1 2 4 2 4 4 1 1 2 4 1 3 1 4 3 5 2 3 0 4 1 2 3 4 4 4 1 4 3 3
                  1 3 3 1 1 4 2 4 1 3 3 1 3 1 3 2 1 2 4 3 0 2 4 2 2 4 2 2 3 4 4 1 4 3 3 4 3 4 4 2 4 1 1 3 2 3 2 4 0 0
                  0 0 0 0 0 0 0 0 0 0

579364919   1     5 5 1 4 2 2 1 2 1 4 2 2 2 2 4 1 3 3 2 2 3 3 1 3 3 5 1 2 3 3 4 3 2 4 2 3 1 2 3 3 3 4 1 1 2 3 3 4 3 2
                  3 3 4 3 3 3 4 3 3 4 2 4 4 4 2 4 1 2 2 4 1 2 3 2 4 3 3 3 3 4 4 5 3 4 3 5 3 4 0 4 2 2 1 4 2 4 1 2 3 1
                  1 4 3 1 1 4 2 4 4 3 1 2 4 1 3 2 2 2 4 4 0 1 2 4 2 1 2 4 2 3 3 1 3 3 3 4 3 1 2 2 2 1 3 4 2 4 4 1 0 0
                  0 0 0 0 0 0 0 0 0 0

577502224   1     5 3 2 2 2 3 1 2 4 5 4 5 5 4 3 4 2 5 1 1 2 4 4 2 2 4 2 2 4 1 2 1 2 3 1 2 1 2 3 2 4 3 4 3 2 3 3 4 4 1
                  3 3 3 3 3 2 4 3 3 4 2 4 2 4 3 4 2 1 1 4 4 3 1 4 4 4 3 1 3 4 4 3 1 3 1 5 2 1 0 1 5 2 1 4 2 1 1 4 3 1
                  2 2 3 1 4 2 2 3 3 3 1 2 3 4 3 4 2 3 4 4 0 1 4 4 1 4 2 3 4 3 2 3 3 1 1 4 3 2 2 1 1 2 1 1 2 1 2 1 0 0
                  0 0 0 0 0 0 0 0 0 0
```

continued

```
   THIS IS A TEST OF  UOM32
1.D.NO                   SCORE  T-SCORE  NO.RIGHT  NO.WRONG  NO.OMIT  CODE

577609528                  134       88       134         0        0     1
119264891                   78       53        78        55        1     1
162329469                   79       53        79        55        0     1
217407036                   66       45        66        68        0     1
578307489                   76       51        76        57        1     1
197305130                   70       48        70        57        7     1
218568291                   85       57        85        49        0     1
219488273                   65       45        65        69        0     1
577623567                   57       40        57        73        4     1
577440908                   65       45        65        69        0     1
217447773                   67       46        67        66        1     1
579488360                   73       50        73        61        0     1
409668407                   72       49        72        61        1     1
484349444                   81       55        81        52        1     1
094320458                   75       51        75        59        0     1
215425147                   62       43        62        71        1     1
219421288                   57       40        57        77        0     1
579384919                   70       48        70        64        0     1
577502224                   68       46        68        66        0     1
```

```
FREQUENCY DISTRIBUTION OF SCORES
ONE * PER STUDENT

   0   0  @
   1   0  @
   2   0  @
   3   0  @
   4   0  @
   5   0  @
   6   0  @
   7   0  @
   8   0  @
   9   0  @
  10   0  @

       .

       .

       .

  45   0  @
  46   0  @
  47   0  @
  48   0  @
  49   0  @
  50   0  @
  51   0  @
```

continued

```
  52   0  @
  53   0  @
  54   0  @
  55   0  @
  56   0  @
  57   2  @**
  58   0  @
  59   0  @
  60   0  @
  61   0  @
  62   1  @*
  63   0  @
  64   0  @
  65   2  @**
  66   1  @*
  67   1  @*
  68   1  @*
  69   0  @
  70   2  @**
  71   0  @
  72   1  @*
  73   1  @*
  74   0  @
  75   1  @*
  76   1  @*
  77   0  @
  78   1  @*
  79   1  @*
  80   0  @
  81   1  @*
  82   0  @
  83   0  @
  84   0  @
  85   1  @*
  86   0  @
  87   0  @
  88   0  @
  89   0  @
      .
      .
      .
 130   0  @
 131   0  @
 132   0  @
 133   0  @
 134   1  @*
 135   0  @
 136   0  @
      .
      .
      .
 155   0  @
 156   0  @
 157   0  @
 158   0  @
 159   0  @
 160   0  @




MEAN OF SCORES = 73.684     S.D. OF SCORES = 16.033
     TOTAL STUDENTS =  19.
```

continued

```
   THIS IS A TEST OF  UOM32
QUESTION   1      1      2      3      4      5  OMIT      CORRECT ANSWER IS  5

     UPPER        0      1      0      0      9     0      P =     .842    PHI =    .167
     LOWER        0      2      0      0      7     0      RPBI=    .136   RB  =    .204     DIS =    .211

QUESTION   2      1      2      3      4      5  OMIT      CORRECT ANSWER IS  3

     UPPER        0      0      5      1      4     0      P =     .526    PHI =   -.056
     LOWER        2      0      5      0      2     0      RPBI=    .238   RB  =    .298     DIS =    .000

QUESTION   3      1      2      3      4      5  OMIT      CORRECT ANSWER IS  1

     UPPER        7      0      0      3      0     0      P =     .526    PHI =    .367
     LOWER        3      2      0      4      0     0      RPBI=    .356   RB  =    .446     DIS =    .421

QUESTION   4      1      2      3      4      5  OMIT      CORRECT ANSWER IS  1

     UPPER        5      1      0      4      0     0      P =     .474    PHI =    .056
     LOWER        4      1      0      4      0     0      RPBI=    .144   RB  =    .180     DIS =    .105

QUESTION   5      1      2      3      4      5  OMIT      CORRECT ANSWER IS  2

     UPPER        0      9      0      0      0     1      P =     .947    PHI =    .000
     LOWER        0      9      0      0      0     0      RPBI=    .054   RB  =    .116     DIS =    .000

                 .                                              .
                 .                                              .
                 .                                              .


QUESTION 156      1      2      3      4      5  OMIT      QUESTION OMITTED

     UPPER        0      0      0      0      0    10      P =     .000    PHI =    .000
     LOWER        0      0      0      0      0     9      RPBI=    .000   RB  =    .000     DIS =    .000

QUESTION 157      1      2      3      4      5  OMIT      QUESTION OMITTED

     UPPER        0      0      0      0      0    10      P =     .000    PHI =    .000
     LOWER        0      0      0      0      0     9      RPBI=    .000   RB  =    .000     DIS =    .000

QUESTION 158      1      2      3      4      5  OMIT      QUESTION OMITTED

     UPPER        0      0      0      0      0    10      P =     .000    PHI =    .000
     LOWER        0      0      0      0      0     9      RPBI=    .000   RB  =    .000     DIS =    .000

QUESTION 159      1      2      3      4      5  OMIT      QUESTION OMITTED

     UPPER        0      0      0      0      0    10      P =     .000    PHI =    .000
     LOWER        0      0      0      0      0     9      RPBI=    .000   RB  =    .000     DIS =    .000

QUESTION 160      1      2      3      4      5  OMIT      QUESTION OMITTED

     UPPER        0      0      0      0      0    10      P =     .000    PHI =    .000
     LOWER        0      0      0      0      0     9      RPBI=    .000   RB  =    .000     DIS =    .000




TEST RELIABILITY K-R(20) =   .907

TEST RELIABILITY K-R(21) =   .851

END OF OUTPUT FOR THIS JOB.
```

COST ESTIMATE

For the program listed on the Sample Input, the running time was 7.092 seconds for the central processor. At the current rate for the University of Maryland's UNIVAC 1108 ($720./hr), the chargeable computer time was $1.42.

Charge to user = computer time + postage and handling + network overhead
= $1.42 + $15.00 + network overhead
= $16.42 + network overhead

CONTENTS

pages
1– 4 Identification & Abstract
5– 7 User Instructions
9–15 I/O
17 Cost—Contents

DESCRIPTIVE TITLE	Spearman Rank-Order Correlation
CALLING NAME	UOM 4
INSTALLATION NAME	University of Maryland
AUTHOR(S) AND AFFILIATION(S)	James A. Forbes University of Maryland
LANGUAGE	FORTRAN II
COMPUTER	IBM 7094
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Sam Steinberg, Prog. Lib., Computer Science Center, Univ. of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This program computes a Spearman rank-order correlation coefficient and a t ratio to determine whether the correlation coefficient is significantly different from zero. Data are read in a two dimensional matrix where there are K variables with N values of each variable. The portion of the correlation matrix below the diagonal and including the diagonal is not computed. Therefore, where there are K variables, (K!/2!(K-2)!) correlation coefficients are computed. For example, if K=5 the correlation coefficients are computed between variables 1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5, 4-5.

Output from this job includes the following. The two variables being correlated will be called X and Y in this description.

—the identification numbers of the two variables being correlated,
—the correction factor for the total sum of squares of each variable, based on the number of sets of ties in each variable and the number of ties in each set of ties, (SUMTX, SUMTY, TIESX, TIESY) ,
—the corrected sum of squares for each variable, (SUMXSQ, SUMYSQ) ,
—sum of the squared differences between the rank value of X and its corresponding rank value for Y, (SUMDSQ),
—spearman rank-order correlation coefficient, (RHO),
—the total number of values of each variable ,
—a t ratio ,
—degrees of freedom for the t test.

continued

The following output is optionally available to supplement the summary output described above.

—listing of raw X values with their corresponding raw Y values,
—listing of ranked raw X values with their corresponding Y values,
—listing of the ranked rank of X values with their corresponding Y values,
—listing of ranked raw Y values with their corresponding rank value for X,
—listing of ranked rank values for Y with their corresponding rank value for X,
—listing of difference values—the difference between a rank value of X and its corresponding rank value for Y.

Limitations per problem

The number of X values must not exceed 100.
The number of Y values must not exceed 100.
The number of variables to be intercorrelated (K) must not exceed 100.
The number of tied values in any given set of ties must not exceed 5.

USER INSTRUCTIONS

The user's input consists of the following cards in the order listed. System cards will be provided by the University of Maryland personnel.

Number Card

This card indicates the number of problems in the job.

Columns	*Contents*
1– 3	Number of problems

Control Card (One card per problem)

This card gives information about the problem.

Columns	*Contents*
1– 3	Number of Variable Format Cards (NVF)
4– 6	Number of Variables to be intercorrelated (K).
9	Output option (PRENT). 3: print all output 1: print summary output
10–12	Number of values of each variable (N).

Variable Format Card (F-format)

Data Input Cards

REFERENCES

Beam, A., *Preliminary Description: Use of IBM Basic Monitor (IBSYS) and Associated Systems*, (Comp. Sci. Cent., Univ. of Maryland, College Park, Md., 1963).

Dixon, W.J., (Ed.), *Biomedical Computer Programs*, (Depart. of Prevent. Med. and Pub. Health, Sch. of Med., Univ. of Calif., Los Angeles, 1964).

Ferguson, G.A., *Statistical Analysis in Psychology and Education*, (McGraw-Hill Book Co., Inc., New York, 1959).

Guilford, J.P., *Fundamental Statistics in Psychology and Education*, (McGraw-Hill Book Co., Inc., New York, 1956).

McNemar, Q., *Psychological Statistics*, (John Wiley & Son, Inc., New York, 1962), 3rd ed.

SAMPLE INPUT

```
  2
  1  4  3 26
(3X,F4.1,2X,F4.1,2X,F4.1,2X,F4.1)
A  10.0  11.4  11.7  10.8  09.3  10.2  12.6  10.8
B  11.5  12.9  14.7  12.7  10.5  11.3  13.7  11.8
C  10.2  12.2  12.7  11.7  10.8  11.9  11.6  11.5
D  10.3  12.7  13.3  12.1  10.0  11.2  12.2  11.1
E  09.7  10.6  10.9  10.4  09.5  09.7  11.2  10.1
F  10.0  12.5  12.5  11.8  09.5  11.4  10.0  10.4
G  10.5  11.2  11.9  11.1  10.0  10.1  11.5  10.5
H  11.8  12.5  12.1  12.1  08.6  11.6  11.5  10.5
I  10.3  10.7  11.2  10.8  09.2  11.0  10.5  10.3
J  10.8  11.1  10.7  10.9  08.8  10.8  10.2  09.9
K  08.1  09.0  09.4  08.9  07.5  08.9  09.1  08.4
L  10.4  11.7  13.6  11.7  09.0  10.6  11.5  10.4
M  11.2  12.5  13.4  12.4  10.5  10.3  12.1  11.0
N  10.2  10.9  11.8  10.8  08.9  10.2  10.9  10.0
O  11.9  10.2  11.9  11.3  09.2  11.4  11.0  10.7
P  11.2  11.2  12.1  11.5  09.9  10.6  12.1  10.7
Q  09.1  10.0  09.9  09.6  09.6  08.8  09.2  09.1
R  10.9  12.9  12.7  12.1  10.6  11.4  11.3  11.1
S  10.8  13.7  12.8  12.5  10.1  11.4  13.0  11.6
T  11.2  12.8  12.6  12.3  10.7  11.4  12.5  11.6
U  08.8  09.9  10.6  09.6  08.9  08.4  10.4  09.1
V  09.6  09.5  09.7  09.6  08.0  09.4  09.3  08.9
W  10.4  12.6  13.0  12.2  10.6  10.6  11.5  10.8
X  08.5  09.2  10.4  09.6  07.6  09.5  08.7  08.5
Y  09.0  10.0  09.7  09.7  06.9  10.4  09.3  09.1
Z  07.3  09.6  09.6  09.0  07.4  08.6  08.0  08.2
  1  8  1 26
(3X,F4.1,2X,F4.1,2X,F4.1,2X,F4.1,2X,F4.1,2X,F4.1,2X,F4.1,2X,F4.1)
```

continued

A	10.0	11.4	11.7	10.8	09.3	10.2	12.6	10.8
B	11.5	12.9	14.7	12.7	10.5	11.3	13.7	11.8
C	10.2	12.2	12.7	11.7	10.8	11.9	11.6	11.5
D	10.3	12.7	13.3	12.1	10.0	11.2	12.2	11.1
E	09.7	10.6	10.9	10.4	09.5	09.7	11.2	10.1
F	10.0	12.5	12.5	11.8	09.5	11.4	10.0	10.4
G	10.5	11.2	11.9	11.1	10.0	10.1	11.5	10.5
H	11.8	12.5	12.1	12.1	08.6	11.6	11.5	10.5
I	10.3	10.7	11.2	10.8	09.2	11.0	10.5	10.2
J	10.8	11.1	10.7	10.9	08.8	10.8	10.2	09.9
K	08.1	09.0	09.4	08.9	07.5	08.9	09.1	08.4
L	10.4	11.7	13.6	11.7	09.0	10.6	11.5	10.4
M	11.2	12.5	13.4	12.4	10.5	10.3	12.1	11.0
N	10.2	10.9	11.8	10.8	08.9	10.2	10.9	10.0
O	11.9	10.2	11.9	11.3	09.2	11.4	11.0	10.7
P	11.2	11.2	12.1	11.5	09.9	10.6	12.1	10.7
Q	09.1	10.0	09.9	09.6	09.6	08.8	09.2	09.1
R	10.9	12.9	12.7	12.1	10.6	11.4	11.3	11.1
S	10.8	13.7	12.8	12.5	10.1	11.4	13.0	11.6
T	11.2	12.8	12.6	12.3	10.7	11.4	12.5	11.6
U	08.8	09.9	10.6	09.6	08.9	08.4	10.4	09.1
V	09.6	09.5	09.7	09.6	08.0	09.4	09.3	08.9
W	10.4	12.6	13.0	12.2	10.6	10.6	11.5	10.8
X	08.5	09.2	10.4	09.6	07.6	09.5	08.7	08.5
Y	09.0	10.0	09.7	09.7	06.9	10.4	09.3	09.1
Z	07.3	09.6	09.6	09.0	07.4	08.6	08.0	08.2

SAMPLE OUTPUT

```
                                        RAW X VALUES WITH RAW Y VALUES
X  10.0   11.5   10.2   10.3    9.7   10.0   10.5   11.8   10.3   10.8    8.1   10.4   11.2   10.2   11.9
X  11.2    9.1   10.9   10.8   11.2    8.8    9.6   10.4    8.5    9.0    7.3
Y  11.4   12.9   12.2   12.7   10.6   12.5   11.2   12.5   10.7   11.1    9.0   11.7   12.5   10.9   10.2
Y  11.2   10.0   12.9   13.7   12.8    9.9    9.5   12.6    9.2   10.0    9.6

                                   -    RANKED RAW X VALUES WITH RAW Y VALUES
X   7.3    8.1    8.5    8.8    9.0    9.1    9.6    9.7   10.0   10.0   10.2   10.2   10.3   10.3   10.4
X  10.4   10.5   10.8   10.8   10.9   11.2   11.2   11.2   11.5   11.8   11.9
Y   9.6    9.0    9.2    9.9   10.0   10.0    9.5   10.6   12.5   11.4   12.2   10.9   12.7   10.7   12.6
Y  11.7   11.2   11.1   13.7   12.9   11.2   12.5   12.8   12.9   12.5   10.2

                                          RANKED RANK OF X WITH RAW Y VALUES
X   1.0    2.0    3.0    4.0    5.0    6.0    7.0    8.0    9.5    9.5   11.5   11.5   13.5   13.5   15.5
X  15.5   17.0   18.5   18.5   20.0   22.0   22.0   22.0   24.0   25.0   26.0
Y   9.6    9.0    9.2    9.9   10.0   10.0    9.5   10.6   12.5   11.4   12.2   10.9   12.7   10.7   12.6
Y  11.7   11.2   11.1   13.7   12.9   11.2   12.5   12.8   12.9   12.5   10.2

                                          RANK OF X WITH RANKED RAW Y VALUES
X   2.0    3.0    7.0    1.0    4.0    6.0    5.0   26.0    8.0   13.5   11.5   18.5   22.0   17.0    9.5
X  15.5   11.5   22.0   25.0    9.5   15.5   13.5   22.0   20.0   24.0   18.5
Y   9.0    9.2    9.5    9.6    9.9   10.0   10.0   10.2   10.6   10.7   10.9   11.1   11.2   11.2   11.4
Y  11.7   12.2   12.5   12.5   12.5   12.6   12.7   12.8   12.9   12.9   13.7

                                       RANK OF X WITH RANKED RANK OF Y
X   2.0    3.0    7.0    1.0    4.0    6.0    5.0   26.0    8.0   13.5   11.5   18.5   22.0   17.0    9.5
X  15.5   11.5   22.0   25.0    9.5   15.5   13.5   22.0   20.0   24.0   18.5
Y   1.0    2.0    3.0    4.0    5.0    6.5    6.5    8.0    9.0   10.0   11.0   12.0   13.5   13.5   15.0
Y  16.0   17.0   19.0   19.0   19.0   21.0   22.0   23.0   24.5   24.5   26.0

                                       DIFFERENCE VALUES
    1.0    1.0    4.0   -3.0   -1.0    -.5   -1.5   18.0   -1.0    3.5     .5    6.5    8.5    3.5   -5.5
    -.5   -5.5    3.0    6.0   -9.5   -5.5   -8.5   -1.0   -4.5    -.5   -7.5

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ   SUMYXQ   SUMDSQ      RHO  TOTAL N  T RATIO       DF
  1-  2       4.50     6.00     3.50     4.00  1458.00  1459.00   871.00      .70    26.00     4.82    24.00
```

continued

```
                                         RAW X VALUES WITH RAW Y VALUES
X
  11.7   14.7   12.7   13.3   10.9   12.5   11.9   12.1   11.2   10.7    9.4   13.6   13.4   11.8   11.9
X
  12.1    9.9   12.7   12.8   12.6   10.6    9.7   13.0   10.4    9.7    9.6
Y
  10.8   12.7   11.7   12.1   10.4   11.8   11.1   12.1   10.8   10.9    8.9   11.7   12.4   10.8   11.3
Y
  11.5    9.6   12.1   12.5   12.3    9.6    9.6   12.2    9.6    9.7    9.0

                                         RANKED RAW X VALUES WITH RAW Y VALUES
X
   9.4    9.6    9.7    9.7    9.9   10.4   10.6   10.7   10.9   11.2   11.7   11.8   11.9   11.9   12.1
X
  12.1   12.5   12.6   12.7   12.7   12.8   13.0   13.3   13.4   13.6   14.7
Y
   8.9    9.0    9.7    9.6    9.6    9.6    9.6   10.9   10.4   10.8   10.8   10.8   11.1   11.3   12.1
Y
  11.5   11.8   12.3   11.7   12.1   12.5   12.2   12.1   12.4   11.7   12.7

                                         RANKED RANK OF X WITH RAW Y VALUES
X
   1.0    2.0    3.5    3.5    5.0    6.0    7.0    8.0    9.0   10.0   11.0   12.0   13.5   13.5   15.5
X
  15.5   17.0   18.0   19.5   19.5   21.0   22.0   23.0   24.0   25.0   26.0
Y
   8.9    9.0    9.7    9.6    9.6    9.6    9.6   10.9   10.4   10.8   10.8   10.8   11.1   11.3   12.1
Y
  11.5   11.8   12.3   11.7   12.1   12.5   12.2   12.1   12.4   11.7   12.7

                                         RANK OF X WITH RANKED RAW Y VALUES
X
   1.0    2.0    3.5    5.0    6.0    7.0    3.5    9.0   10.0   11.0   12.0    8.0   13.5   13.5   15.5
X
  19.5   25.0   17.0   19.5   23.0   15.5   22.0   18.0   24.0   21.0   26.0
Y
   8.9    9.0    9.6    9.6    9.6    9.6    9.7   10.4   10.8   10.8   10.8   10.9   11.1   11.3   11.5
Y
  11.7   11.7   11.8   12.1   12.1   12.1   12.2   12.3   12.4   12.5   12.7

                                         RANK OF X WITH RANKED RANK OF Y
X
   1.0    2.0    3.5    5.0    6.0    7.0    3.5    9.0   10.0   11.0   12.0    8.0   13.5   13.5   15.5
X
  19.5   25.0   17.0   19.5   23.0   15.5   22.0   18.0   24.0   21.0   26.0
Y
   1.0    2.0    4.5    4.5    4.5    4.5    7.0    8.0   10.0   10.0   10.0   12.0   13.0   14.0   15.0
Y
  16.5   16.5   18.0   20.0   20.0   20.0   22.0   23.0   24.0   25.0   26.0
-------------------------------------------------------------------------------------------------------
                                         DIFFERENCE VALUES
    .0     .0   -1.0     .5    1.5    2.5   -3.5    1.0     .0    1.0    2.0   -4.0     .5    -.5     .5
   3.0    8.5   -1.0    -.5    3.0   -4.5     .0   -5.0     .0   -4.0     .0

VARIABLES    SUMTX    TIESX    SUMTY    TIESY    SUMXSQ    SUMYXQ    SUMDSQ    RHO   TOTAL N   T RATIO     DF
 3-  4        2.00     4.00     9.50     4.00   1460.50   1453.00    197.50    .93    26.00     12.62     24.00
```

continued

```
VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  1- 2        4.50     6.00     3.50     4.00  1458.00   1459.00    871.00      .70    26.00      4,82    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  1- 3        4.50     6.00     2.00     4.00  1458.00   1460.50    836.00      .71    26.00      4,99    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  1- 4        4.50     6.00     9.50     4.00  1458.00   1453.00    542.00      .81    26.00      6,86    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  1- 5        4.50     6.00     3.00     6.00  1458.00   1459.50   1326.00      .55    26.00      3,19    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  1- 6        4.50     6.00    12.50     3.00  1458.00   1450.00    823.00      .72    26.00      5,04    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  1- 7        4.50     6.00     6.00     3.00  1458.00   1456.50    884.50      .70    26.00      4,76    24.00

                         .                                .                                      .
                         .                                .                                      .
                         .                                .                                      .

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  4- 7        9.50     4.00     6.00     3.00  1453.00   1456.50    541.50      .81    26.00      6,86    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  4- 8        9.50     4.00     5.00     7.00  1453.00   1457.50    269.00      .91    26.00     10,59    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  5- 6        3.00     6.00    12.50     3.00  1459.50   1450.00   1417.00      .51    26.00      2,93    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  5- 7        3.00     6.00     6.00     3.00  1459.50   1456.50    778.50      .73    26.00      5,28    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  5- 8        3.00     6.00     5.00     7.00  1459.50   1457.50    401.00      .86    26.00      8,35    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  6- 7       12.50     3.00     6.00     3.00  1450.00   1456.50   1289.50      .56    26.00      3,28    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  6- 8       12.50     3.00     5.00     7.00  1450.00   1457.50    741.00      .75    26.00      5,47    24.00

VARIABLES    SUMTX    TIESX    SUMTY    TIESY   SUMXSQ    SUMYXQ    SUMDSQ      RHO  TOTAL N   T RATIO      DF
  7- 8        6.00     3.00     5.00     7.00  1456.50   1457.50    231.50      .92    26.00     11,55    24.00
```

COST ESTIMATE

For the job listed on the Sample Input, the running time was 12.532 seconds in the system. At the current rate for the University of Maryland's 7094 ($222./hr.), the chargeable computer time was $0.77.

Charge to user = computer time + postage and handling + network overhead
= $0.77 + $15.00 + network overhead
= $15.77 + network overhead

CONTENTS—UOM 4

pages
1– 2 Identification & Abstract
3 User Instructions
5– 9 I/O
11 Cost—Contents

200

DESCRIPTIVE TITLE	Weighted Summing Test Scoring Program, (Recoding)
CALLING NAME	UOM 87
INSTALLATION NAME	University of Maryland
AUTHOR(S) AND AFFILIATION(S)	A.F. Norcio University of Maryland
LANGUAGE	FORTRAN
COMPUTER	UNIVAC 1108
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Sam Steinberg, Prog. Lib., Computer Science Center, Univ. of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This program selects those variables specified by the user, gives the values of those variables the weights specified by the user and adds the weighted values to a sum. Data may be read from cards or tape. The user is cautioned that the program performs operations that are frequently called recoding and not weighted summation. There is, therefore, a potential semantic difficulty.

Limitations

The user may submit as many separate problems as he desires through the program in only one run. Within each problem the program can handle a maximum of 9,999 observations; for each observation the number of variables must not exceed 999. For each observation the program can compile a maximum of 99 separate sums (scores).

The user may specify up to 1,782 variables for selection; and, it should be noted that any variable may be selected more than once; in addition, a maximum of 891 weights may also be specified.

For each sum (score) that the user desires there must be one Score Card. The user must also specify the format of his data on Variable Format Cards.

continued

Summary of Limitations

a) Number of Problems = unlimited
b) Maximum number of Observations = 9,999
c) Maximum number of Variables = 999
d) Maximum number of Weights = 891
e) Maximum number of Selections = 1,782
f) Maximum number of Sums = 99
g) Maximum number of Variable Format Cards = 10

USER INSTRUCTIONS

The job deck consists of the following cards. Information should be right-justified in the specified fields. Any columns not mentioned should be left blank.

Problem Card

Each problem must begin with a PROBLM card; this card designates that a problem is to be stated, together with the specifications of the problem.

Columns	*Contents*
1– 6	PROBLM (Mandatory)
8–13	Alphanumeric problem code
15–17	Number of variables (not to exceed 999)
19–22	Number of observations (not to exceed 9,999)
24–25	Number of Item Selection Cards (not to exceed 99)
27–28	Number of Weight Cards (not to exceed 99)
30–31	Number of input tape, if not 5
33–34	Number of output tape, if not 6
36–37	Number of Score Cards (not to exceed 99)
39	Output Code 1: punched and printed output 2: printed output only
71–72	Number of Variable Format Cards (not to exceed 10)

Item Selection Cards

On each Item Selection card the user specifies those items that are to be selected for computation of a particular sum.

Columns	*Contents*
1– 6	ITEMSL (Mandatory)
8–13	Alphanumeric identification code
16–17	Number of items that appear on this card (not to exceed 18)
19–21	First item selected
22–24	Section item selected
⋮	⋮
70–72	Eighteenth item selected

continued

Note: There may be a maximum of 99 item selection cards. Alphanumeric codes used on these cards must agree with those codes that appear on the Score Cards.

Weight Cards

On these cards the user specifies the weights that he wishes to assign to the various original values of his data.

Columns	*Contents*
1– 6	WEIGHT (Mandatory)
8–13	Alphanumeric identification code
16–17	*Total* number of values that appear on this card
19–21	First original value
22–24	Weight to be assigned to first original value
25–27	Second original value
28–30	Weight to be assigned to second original value
⋮	⋮
67–69	Ninth original value
70–72	Weight to be assigned to ninth original value

Note: There may be a maximum of 99 weight cards. Alphanumeric codes that appear on these cards must agree with those codes that appear on Score Cards.

Score Cards

On these cards the user specifies by alphanumeric codes those combinations of items selected and weights which he desires to be used for computing the various sums (scores).

Columns	*Contents*
1– 6	SCORES (Mandatory)
8–13	Alphanumeric code of Score (this code is used to label the output with the name of this particular score)
16–17	Number of Alphanumeric codes of ITEMSL and WEIGHT cards that appear on this Score Card
19–24	ITEMSL Code

continued

Columns	*Contents*
25–30	WEIGHT Code for these items
31	Blank
32–37	ITEMSL Code
38–43	WEIGHT Code for these items
44	Blank
45–50	ITEMSL Code
51–56	WEIGHT Code for these items
57	Blank
58–63	ITEMSL Code
64–69	WEIGHT Code for these items

Note: For each score desired there must be one Score Card; there may be up to a maximum of 99 Score Cards. Alphanumeric codes that appear on these cards must agree with those codes that appear on ITEMSL and WEIGHT cards.

Variable Format Cards

For preparation of these cards the user is referred to the BMD Manual[1]. There may be up to a maximum of 10 Variable Format Cards.

Data Deck (if data come from cards)

Finish Card

This card is placed after the Data Deck of the last problem; that is, it must physically be the last card that appears in the deck set-up. It designates that the program has completed the final problem.

Columns	*Contents*
1– 6	FINISH (Mandatory)

Deck Set-up

PROBLM Card
ITEMSL Cards (not to exceed 99)
WEIGHT Cards (not to exceed 99)
SCORE Cards (not to exceed 99)
Variable Format Cards
Data Deck

may be repeated as often as desired

:

FINISH Card

continued

Input and Output of Data

The program is able to read data from a user supplied data tape; if this is desired, the logical tape number must be specified in Cols. 30–31 on the PROBLM card; if the data are to be read from cards, these columns should be left blank. The user may also have his output written on tape; if this is desired, he must specify the logical tape number in Cols. 33–34 on the PROBLM card; otherwise, the columns should also be left blank. In addition, if punched and printed output is desired, a 1 must appear in Col. 39 of the PROBLM card; if only printed output is desired, a 2 must appear in this column.

REFERENCE

1. Dixon, W.J., (Ed.), *Biomedical Computer Programs*, (Dept. of Prevent. Med. and Pub. Health, Sch. of Med., Univ. of Calif., Los Angeles, 1964), Section III, pp. 22–28.

SAMPLE INPUT

```
PROBLM  RUN 3 104    20 15  3          6 2                          1
ITEMSL      F1  10   1 11 13 16 22 24 36 46 54 60
ITEMSL      F2  10   2  6 18 28 31 34 38 44 45 51
ITEMSL     SD1   8   3  7 19 21 27 37 56 58
ITEMSL     SD2   9   9 14 23 33 40 42 47 50 63
ITEMSL     SA1   9   8 29 30 41 48 52 57 61 64
ITEMSL     SA2   8   5 12 17 26 35 43 53 59
ITEMSL     OC1   5   4 10 15 20 39
ITEMSL     OC2   5  25 32 49 55 62
ITEMSL      T1  18   1  3  4  7  8 10 11 13 15 16 19 20 21 22 24 27 29 30
ITEMSL      T2  14  36 37 39 41 46 48 52 54 56 57 58 60 61 64
ITEMSL      T3  18   2  5  6  9 12 14 17 18 23 25 26 28 31 32 33 34 35 38
ITEMSL      T4  14  40 42 43 44 45 47 49 50 51 53 55 59 62 63
ITEMSL      D1  18  65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82
ITEMSL      D2  18  83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99100
ITEMSL      D3   4 101102103104
WEIGHT    WTS1  12   1 -3  2 -2  3 -1  4  1  5  2  6  3
WEIGHT    WTS2  12   1  3  2  2  3  1  4 -1  5 -2  6 -3
WEIGHT     WTD  12   1  1  2  2  3  3  4  5  5  6  6  7
SCORES F SCLE    4      F1  WTS1      F2  WTS2
SCORES SD SCL    4     SD1  WTS1     SD2  WTS2
SCORES SA SCL    4     SA1  WTS1     SA2  WTS2
SCORES OC SCL    4     OC1  WTS1     OC2  WTS2
SCORES  TOTAL    8      T1  WTS1      T2  WTS1      T3  WTS2      T4  WTS2
SCORES  D SCL    6      D1   WTD      D2   WTD      D3   WTD
(1X,A6,64F1.0/7X,40F1.0)
11216A65452543545322425253444432434452343435444434452454544322355435444
21216A6334542353434444453434233222224323422223
11220A64241444325744463455444422463532343254254344452356523435632325232
                  .                                   .

                  .                                   .

                  .                                   .

21210A-116651142621113554211511511532416631211
11214A65441444124112424145454263254554325436443211661256255231654215543
21214A-1354 21136422147541252311443113422411221
10616A65151166551412312134452131651245616313321233221151116563325411423
20616A51213111115121211211641
PROBLM  RUN 4 104    14 15  3          6 2                          1
ITEMSL      F1  10   1 11 13 16 22 24 36 46 54 60
ITEMSL      F2  10   2  6 18 28 31 34 38 44 45 51
ITEMSL     SD1   8   3  7 19 21 27 37 56 58
ITEMSL     SD2   9   9 14 23 33 40 42 47 50 63
ITEMSL     SA1   9   8 29 30 41 48 52 57 61 64
ITEMSL     SA2   8   5 12 17 26 35 43 53 59
ITEMSL     OC1   5   4 10 15 20 39
ITEMSL     OC2   5  25 32 49 55 62
ITEMSL      T1  18   1  3  4  7  8 10 11 13 15 16 19 20 21 22 24 27 29 30
ITEMSL      T2  14  36 37 39 41 46 48 52 54 56 57 58 60 61 64
ITEMSL      T3  18   2  5  6  9 12 14 17 18 23 25 26 28 31 32 33 34 35 38
ITEMSL      T4  14  40 42 43 44 45 47 49 50 51 53 55 59 62 63
ITEMSL      D1  18  65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82
ITEMSL      D2  18  83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99100
ITEMSL      D3   4 101102103104
WEIGHT    WTS1  12   1 -3  2 -2  3 -1  4  1  5  2  6  3
WEIGHT    WTS2  12   1  3  2  2  3  1  4 -1  5 -2  6 -3
WEIGHT     WTD  12   1  1  2  2  3  3  4  5  5  6  6  7
SCORES F SCLE    4      F1  WTS1      F2  WTS2
SCORES SD SCL    4     SD1  WTS1     SD2  WTS2
SCORES SA SCL    4     SA1  WTS1     SA2  WTS2
SCORES OC SCL    4     OC1  WTS1     OC2  WTS2
SCORES  TOTAL    8      T1  WTS1      T2  WTS1      T3  WTS2      T4  WTS2
SCORES  D SCL    6      D1   WTD      D2   WTD      D3   WTD
(1X,A6,64F1.0/7X,40F1.0)
11201A7216112664511225133265156164567444256151611443264245124666166324
21201A74234213144263414312222422423423553344 1221
11206A711526426651114566026555145264661443416141341661144665134656146626
21206A715152141463342134111416312232125245511111
                  .                                   .

                  .                                   .

                  .                                   .

21214A72443231335343451522452222322222323332-33
11210A73451116653111143165432153163461131636151111444166155155654115414
21210A71146611656242111654221123331533214651-21
11221A74141344613414241144443144244446344414314441661344664644114415666
21221A71116134444664311463331331633214344431643
FINISH
```

SAMPLE OUTPUT

```
UOM87    TEST SCORING

PROGRAM LIBRARY
COMPUTER SCIENCE CENTER
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND

VERSION OF 6/1/67

PROGRAMMED BY A. FRANCIS NORCIO



PROBLEM CODE= RUN 3
NUMBER OF OBSERVATIONS=   20
NUMBER OF VARIABLES= 104
NUMBER OF ITEM SELECTION CARDS= 15
NUMBER OF WEIGHT CARDS=  3
NUMBER OF DATA INPUT TAPE= 5
NUMBER OF OUTPUT TAPE= 6
NUMBER OF VARIABLE FORMAT CARDS=  1

ITEMS SELECTED FOR      F1
        1
       11
       13
       16
       22
       24
       36
       46
       54
       60

ITEMS SELECTED FOR      F2
        2
        6
       18
       28
       31
       34
       38
       44
       45
       51

                 .

                 .

                 .

ITEMS SELECTED FOR      C3
      101
      102
      103
      104

                   WTS1
ORIGINAL VALUES             WEIGHTS
        1.                    -3.
        2.                    -2.
        3.                    -1.
        4.                     1.
        5.                     2.
        6.                     3.

                   WTS2
ORIGINAL VALUES             WEIGHTS
        1.                     3.
        2.                     2.
        3.                     1.
        4.                    -1.
        5.                    -2.
        6.                    -3.
```

continued

```
                  WTD

ORIGINAL VALUES           WEIGHTS
      1.                     1.
      2.                     2.
      3.                     3.
      4.                     5.
      5.                     6.
      6.                     7.

TOTAL NUMBER OF ITEMS SELECTED= 168
TOTAL NUMBER OF WEIGHTS=  18
TOTAL NUMBER OF SETS=  15

VARIABLE FORMAT IS
(1X,A6,64F1.0/7X,40F1.0)


   1216A6 F SCLE -11.00 SD SCL  -5.00 SA SCL  13.00 OC SCL   3.00  TOTAL  -0.
          D SCL 140.00
   1220A6 F SCLE   2.00 SD SCL   7.00 SA SCL  -4.00 OC SCL   5.00  TOTAL  10.00
          D SCL 134.00
   1215A6 F SCLE  -1.00 SD SCL  18.00 SA SCL  32.00 OC SCL  -3.00  TOTAL  46.00
          D SCL 145.00
   1202A6 F SCLE  -4.00 SD SCL   7.00 SA SCL  20.00 OC SCL  -0.     TOTAL  23.00
          D SCL 153.00
                                        .                  .
               .                        .                  .
               .                        .                  .
               .


   1214A6 F SCLE -22.00 SD SCL  27.00 SA SCL  19.00 OC SCL  -1.00  TOTAL  23.00
          D SCL 115.00
   C616A6 F SCLE  -3.00 SD SCL  28.00 SA SCL  -1.00 OC SCL -12.00  TOTAL  12.00
          D SCL  43.00




UOM      TEST SCORING

PROGRAM LIBRARY
COMPUTER SCIENCE CENTER
UNIVERSITY OF MARYLAND
COLLEGE PARK, MARYLAND

VERSION OF 6/1/67

PROGRAMMED BY A. FRANCIS NORCIO


PROBLEM CODE= RUN 4
NUMBER OF OBSERVATIONS=   14
NUMBER OF VARIABLES= 104
NUMBER OF ITEM SELECTION CARDS= 15
NUMBER OF WEIGHT CARDS=  3
NUMBER OF DATA INPUT TAPE= 5
NUMBER OF OUTPUT TAPE= 6
NUMBER OF VARIABLE FORMAT CARDS=  1

ITEMS SELECTED FOR      F1
        1
       11
       13
       16
       22
       24
       36
       46
       54
       60
```

continued

```
ITEMS SELECTED FOR      F2
        2
        6
       18
       28
       31
       34
       38
       44
       45
       51

                    .
                    .
                    .
                    .

ITEMS SELECTED FOR      C3
      101
      102
      103
      104

                         WTS1
ORIGINAL VALUES               WEIGHTS
       1.                       -3.
       2.                       -2.
       3.                       -1.
       4.                        1.
       5.                        2.
       6.                        3.

                         WTS2
ORIGINAL VALUES               WEIGHTS
       1.                        3.
       2.                        2.
       3.                        1.
       4.                       -1.
       5.                       -2.
       6.                       -3.

                         WTC


ORIGINAL VALUES               WEIGHTS
       1.                        1.
       2.                        2.
       3.                        3.
       4.                        5.
       5.                        6.
       6.                        7.

TOTAL NUMBER OF ITEMS SELFCTED= 168
TOTAL NUMPER OF WEIGHTS=  18
TOTAL NUMRER OF SETS=  15

VARIABLE FORMAT IS
(1X,A6,64F1.0/7X,40F1.0)


   1201A7 F SCLE   2.00 SD SCL  38.C0 SA SCL  46.00 OC SCL  -4.00  TOTAL  82.00
          D SCL 125.00
   1206A7 F SCLE  -7.00 SD SCL  19.C0 SA SCL  34.00 OC SCL  -2.00  TOTAL  44.00
          D SCL 120.00

              .                  .                  .
              .                  .                  .
              .                  .                  .


   1210A7 F SCLE  -1.00 SD SCL  36.C0 SA SCL  32.00 OC SCL  -2.00  TOTAL  65.00
          D SCL 134.00
   1221A7 F SCLE -16.00 SD SCL   3.C0 SA SCL  17.00 OC SCL -10.00  TOTAL  -6.00
          D SCL 146.00
```

COST ESTIMATE

For the job listed on the Sample Input, the running time was 0.77 seconds for the central processor. At the current rate for the University of Maryland's UNIVAC 1108 ($720./hr., 2-second minimum), the chargeable computer time was $0.40.

Charge to user = computer time + postage and handling + network overhead
= $0.40 + $15.00 + network overhead
= $15.40 + network overhead

CONTENTS—UOM 87

pages

1– 2 Identification & Abstract
3– 6 User Instructions
7–10 I/O
11 Cost—Contents

DESCRIPTIVE TITLE	Generalized t Test
CALLING NAME	UOM 5
INSTALLATION NAME	University of Maryland Computer Science Center
AUTHOR(S) AND AFFILIATION(S)	J.F. Williams Computer Science Center University of Maryland
LANGUAGE	FORTRAN IV
COMPUTER	IBM 7090/7094 and UNIVAC 1108
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mr. Sam Steinberg, Program Librarian, Computer Science Center, University of Maryland, College Park, Md. 20740 Tel.: (301) 454-4261

FUNCTIONAL ABSTRACT

This program computes the t statistic for differences between the means of independent (random) groups or dependent (matched) groups. The user may select to input his data in either of two ways; by rows, where each card represents one variable in all its observations, or by columns, where each card represents one observation with only the variables included in that observation being punched.

Output includes the following.

—sum of values (SUM) for a given variable or group,

—sum of squared values (SUMSQ) for a given variable or group,

—mean of values (XMEAN) for a given variable or group,

—variance of values (SAMVAR) for a given variable or group,

—standard deviation of values (SAMSD) for a given variable or group,

—t (TEE) for a given pair of variables or groups

—sum of cross-products (SUMCP) and product-moment correlation (R) for a given pair of variables or groups *with a dependent t only*,

continued

—F-ratio (FRATIO) for a given pair of variables or groups *with independent t only.*

Limitations

50 variables of groups (M)
250 observations per variable (N)
3 Variable Format Cards

Computational Formulas

Product-moment correlation coefficient

$$r_{ab} = \sqrt{\frac{[((\Sigma\ X_a X_b/N) - (\Sigma X_a \Sigma X_b))/N^2]^2}{s_a^2\ s_b^2}}$$

Dependent t (using difference scores)

$$t_{ab} = \frac{\overline{X}d}{s_d\sqrt{N}}$$

Independent t

$$t_{ab} = \sqrt{\frac{(\overline{X}_a - \overline{X}_b)^2}{[(N_a s_a^2 + N_b s_b^2) / (N_a + N_b - 2)][(N_a + N_b)/(N_a N_b)]}}$$

REFERENCE

McNemar, Q., *Psychological Statistics*, (John Wiley and Sons, Inc., New York, 1962), 3rd ed.

USER INSTRUCTIONS

Data Input

In using this program, the user may punch the data in one of two ways, i.e., by rows or by columns.

By Rows: each card represents *one* variable or group and all observations for this variable are punched on the card (overflow onto other cards is permitted). T is computed between cards.

By Columns: each card represents *one* observation. All variables or groups for this observation are punched on a card (overflow onto other cards is permitted). T is computed between columns.

Card Preparation

Problem Card

Columns	*Contents*
1– 6	Problem number (may be alphabetical)
7– 9	Number of variables or groups (M)
10–12	Number of Variable Format Cards
13–15	Sample size for the first variable or group
16–18	Sample size for the second variable or group
⋮	⋮
70–72	Sample size for the 20th variable or group

When a second or third card is required, information should be continued in Col. 1 of the new card.

In the three columns immediately after the sample size for the last group the Data Specification should be punched

001: data are punched by columns
002: data are punched by rows

The three columns immediately after the Data Specification, the Variable Type should be punched

001: dependent t is desired
002: independent t is desired

continued

Variable Format Cards

Up to 3 Format Cards with F-type specification are allowed.

Data Cards

Punch the data according to the input option and format specified.

Job Deck

System Cards
Program Deck
Problem Card
Variable Format Cards
Data Cards

(Problem Card, Variable Format Cards, Data Cards) may be repeated for multiple cases

SAMPLE INPUT

```
KITTYD  2  1 24 24001001
(2F2.0)
 4 3
 6 3
 6 0
 4 2
 4 2
 5 3
 4 3
 2 0
 5 2
 5 1
 5 4
 5 2
 5 1
 5 0
 5 0
 3 0
 4 1
 2 0
 6 6
 5 2
 4 2
 5 3
 3 0
 5 2
BIRCON  2  1 14 13001002
(2F3.0)
  3  8
 -2 11
  0  4
-10-25
- 3- 6
  2- 4
  0  8
- 2- 1
  0  2
  2- 1
  8- 4
  2 16
  7  5
  8
CONBIR  2  1 11 13001002
(2F3.0)
 10  8
- 2 11
 10  4
  6-25
 12- 6
- 6- 4
  2  8
- 9- 1
  2  2
  6- 1
  0- 4
    16
     5
```

SAMPLE OUTPUT

```
CALCULATED TEE



PROBLEM CARD


PROBLEM CODE              KITTYD
NUMBER OF TREATMENTS           2
OPTION FOR INPUT DATA          1
OPTION FOR TYPE OF TEE TEST    1
NUMBER OF SUBJECTS PER TREATMENT
  24  24



SUMS OF OBSERVATIONS


   107.00000     42.00000

SUMS OF SQUARES OF OBSERVATIONS


     505.00      128.00

MEANS OF OBSERVATIONS


     4.45833      1.75000

VARIANCES OF OBSERVATIONS


     1.16493      2.27083

STANDARD DEVIATIONS OF OBSERVATIONS


     1.07932      1.50693



SUM OF CROSS PRODUCTS OF OBSERVATIONS



ROW   1
     505.00      206.00

ROW   2
     206.00      128.00



PRODUCT-MOMENT CORRELATIONS BETWEEN TREATMENTS



ROW   1
     1.00000       .48034

ROW   2
      .48034      1.00000

CALCULATED TEE FOR MEAN PAIRS



ROW   1
      .00000      9.49001

ROW   2
    -9.49001       .00000
```

continued

```
CALCULATED TEE



PROBLEM CARD


PROBLEM CODE              BIRCON
NUMBER OF TREATMENTS           2
OPTION FOR INPUT DATA          1
OPTION FOR TYPE OF TEE TEST    2
NUMBER OF SUBJECTS PER TREATMENT
  14  13



SUMS OF OBSERVATIONS

     15.00000    13.00000
SUMS OF SQUARES OF OBSERVATIONS

     315.00     1245.00
MEANS OF OBSERVATIONS

     1.07143      1.00000
VARIANCES OF OBSERVATIONS

     21.35204     94.76923
STANDARD DEVIATIONS OF OBSERVATIONS

     4.62083      9.73495



CALCULATED TEE FOR MEAN PAIRS


ROW   1
     .00000        .02370

ROW   2
     .02370        .00000



ESTIMATES OF POPULATION VARIANCE

     22.99451    102.66667
F RATIO FOR MEAN PAIRS AT N(X)-1 AND N(Y)-1 DF


ROW   1
     1.00000       4.46483

ROW   2
     4.46483       1.00000
```

continued

```
CALCULATED TEE



PROBLEM CARD


PROBLEM CODE              CONAIR
NUMBER OF TREATMENTS           2
OPTION FOR INPUT DATA          1
OPTION FOR TYPE OF TEE TEST    2
NUMBER OF SUBJECTS PER TREATMENT
  11  13



SUMS OF OBSERVATIONS

     31.00000     13.00000

SUMS OF SQUARES OF OBSERVATIONS

      545.00     1245.00

MEANS OF OBSERVATIONS

      2.81818       1.00000

VARIANCES OF OBSERVATIONS

     41.60331     94.76923

STANDARD DEVIATIONS OF OBSERVATIONS

      6.45006      9.73495


CALCULATED TEE FOR MEAN PAIRS


ROW    1
       .00000        .50642

ROW    2
       .50642        .00000



ESTIMATES OF POPULATION VARIANCE

     45.76364    102.66667

F RATIO FOR MEAN PAIRS AT N(X)-1 AND N(Y)-1 DF


ROW    1
      1.00000       2.24341

ROW    2
      2.24341       1.00000
```

COST ESTIMATE

For the job listed on the Sample Input, the computer time was 6.361 seconds. At the current rate for the University of Maryland IBM 7094 ($222./hr.) the chargeable computer time was $0.39.

Charge to user = computer time + postage and handling + network overhead

= $0.39 + $15.00 + network overhead

= $15.39 + network overhead

The same job run on the UNIVAC 1108 used 0.5 sec. of central processor unit (CPU) time. At the current rate for the 1108 ($0.20/sec. with 2 sec. minimum) the computer time cost $0.40.

CONTENTS—UOM 5

pages

1– 2 Identification & Abstract
3– 4 User Instructions
5– 8 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Information Retrieval Services
CALLING NAME	
INSTALLATION NAME	University of Georgia Computer Center
AUTHOR(S) AND AFFILIATION(S)	Information Science Group University of Georgia
LANGUAGE	Assembler Language & PL/1
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution
CONTACT	Miss Margaret Park, Supervisory Inform. Scientist, Computer Center, The Univ. of Ga., Athens, Ga. 30601 Tel.: (404) 542-3741

FUNCTIONAL ABSTRACT

In any endeavor, scientists need to keep constantly abreast of activities in their field of interest, to be on the lookout for new ideas, and to maintain a library of useful references. The proliferation of new scientific knowledge is rapidly outpacing the capabilities of conventional information-handling and publishing techniques. Scientists are now turning to computer-based methods to help speed and channel the flow of information on a timely basis.

The Computer Center at the University of Georgia is actively engaged in establishing an Information Center. Mechanized data bases from several scientific organizations are presently available and in use at the Center. Subject areas currently represented are biology, biochemistry, nuclear science, and chemistry, including structural data files for chemical compounds. Other tape services in fields such as medicine, engineering, physics, geology, etc., will be added as interst is expressed in these subject areas.

The Computer Center's Information Sciences Unit offers assistance in creating search profiles and current awareness and retrospective searches of the scientific literature.

continued

The staff of the Computer Center includes highly trained professionals in practically every subject discipline. In addition to chemists and a microbiologist in the Information Sciences Unit, the Center also has full time staff with specialties in physics, engineering, statistics, biology, and forestry, with consulting staff available in other disciplines. All staff are also familiar with computer systems and their applications.

The Computer Center is constantly seeking better ways to satisfy the information requirements of the scientific community and we will greatly appreciate any suggestions made in this direction. New services will be added as rapidly as possible to meet newly identified needs and uses.

Persons desiring to use the Information Retrieval Services are directed to the contact person.

USER INSTRUCTIONS

In a computer-based system such as the Information Science Center offers, information selected and prepared by scientists is encoded into a machine-readable form. The computer-based record is backed by a complete system of programs for searching and retrieving this information. These search programs are based on word or word-fragment matching, with provisions for term weighting and AND, OR, and NOT logic to combine terms.

The user may specify the data base against which he wishes to have his questions searched, or he may leave the decision up to the Information Scientists who know the subject coverage of the data bases. The University generally makes this decision in consultation with the user, however. The University of Georgia staff does include Information Scientists familiar with both the retrieval techniques as well as the subject disciplines, and these people do all necessary construction and coding of the search profile. The files each vary as to the data content (e.g., titles, abstracts, keywords, descriptors, etc.), the frequency of publication (weekly, biweekly, monthly, quarterly, etc.), subject scope, and types of primary documents covered (e.g., journals, patents, reports, books, etc.). They also vary considerably in the number of years of literature which they cover, some including documents as far back as 1956 while others began only in 1969 or 1970. Due to copyright considerations, each data base is searched independently at the present time, and consequently is priced individually.

Since the University of Georgia staff performs the construction and coding of the search profiles, details of coding are not listed here. The Sample Input and Output show various logical and weighting possibilities in the search programs.

Data bases currently include the following.

Chemical-Biological Activities (*CBAC*) provides up-to-date coverage of published work concerned with the interaction of organic compounds (drugs, pesticides, etc.) with biological systems (man and other plants and animals). Also covered are metabolism studies and studies of in-vitro chemical reactions of biochemical interests. Nearly 700 journals are monitored. *CBAC* is a consolidated source of information in the chemical-biological field, and especially appeals to scientists involved in biochemical and physiological research.

continued

The tape version of *CBAC* corresponds to the *Chemical-Biological Activities* hard-copy version; however, tape issues are usually available two-four weeks ahead of the printed copy. Each issue contains 500-600 digests.

Chemical Titles (*CT*), which is issued biweekly by Chemical Abstracts Service, contains journal references to approximately 4,500 articles per issue appearing in 650 important U.S. and non-U.S. chemical and chemical engineering journals. Titles that appear in *CT* represent over 65% of the total abstracts that later appear in *Chemical Abstracts*. *Chemical Titles* offers journal references to articles approximately 70 days before their abstracts are published in *Chemical Abstracts*. In many cases titles appear in *Chemical Titles* before the journal containing the article is published. Thus *Chemical Titles* is valuable as an alerting service.

CA-Condensates (*CA-C*) is the computer searchable complement to the printed publication, *Chemical Abstracts* (*CA*), which covers the full range of chemistry, referencing 250,000 articles per year.

CA-Condensates is issued weekly; the content corresponds to an issue of *CA*. The tape version, *CA-Condensates*, precedes the corresponding printed issue of *CA* by several weeks due to the time required to print, bind, and distribute *CA* printed issues.

The abstracts in *CA* and *CA-Condensates* are grouped into five categories: Biochemistry, Organic Chemistry, Macromolecular Chemistry, Applied Chemistry and Chemical Engineering, and Physical and Analytical Chemistry. The first two groupings are published as an odd numbered issue one week, and the last three groupings are published as an even numbered issue the following week. Searches may be limited to odd or even numbered issues if desired.

Nuclear Science Abstracts (*NSA*), published semi-monthly by the U.S. Atomic Energy Commission, provides international coverage of the literature on nuclear science and technology. Publications include the scientific and technical reports of the USAEC and its contractors, other government agencies, universities, industrial and research organizations as well as patents, books and journal literature on a world wide basis.

The subject matter coverage of *NSA* includes nuclear chemistry and physics, instrumentation, reactor technology, engineering, earth science and life science.

continued

The tape version of *Bioresearch Index* (*BIORI*) corresponds in coverage to the printed copy of this BioScience Information Service of *Biological Abstracts*. *BIORI*, published monthly with more than 7,000 titles, includes research reports from symposia, congresses, conferences, reports, bulletins, etc. in which biology is the main emphasis. Biological research papers of a more applied nature are covered, also, in *BIORI*. Coverage is not duplicated in *BA*. The tape version of *BIORI* precedes the printed issue of *BIORI* by several weeks due to the time required to print, bind and distribute *BIORI* hard copy.

The tape version of *Biological Abstracts* (*BA*) corresponds in coverage to the printed copy of this BioScience Information Service (BIOSIS) - more than 7,000 journals per year producing more than 200,000 abstracts per year. *BA* issues, containing more than 5,600 titles, are published semi-monthly and chiefly contain abstracts of research papers in biology published in monthly periodicals. Information covered reflects basic research in biology. The tape version of *BA* precedes the printed issue of *BA* by several weeks due to the time required to print, bind and distribute *BA* hard copy.

Educational Resources Information Center (*ERIC*) is a nationwide information network for acquiring, selecting, abstracting, indexing, storing, retrieving, and disseminating the most significant and timely educational research reports and projects.

The *ERIC* program has two component parts: *Research in Education* (*RIE*), and *Current Index to Journals in Education* (*CIJE*).

RIE is a monthly abstract journal announcing recently completed research and research-related projects in the field of education.

CIJE is devoted exclusively to periodical literature, providing detailed indexing for articles in over 500 educational journals. *CIJE* is a monthly companion to *RIE*.

Early in 1970 several additional tape services were added. Data bases added include COMPENDEX from Engineering Index, the USGRDR tapes covering the government reports issued by U.S. Clearinghouse, Bibliography of North American Geology and Selected Water Resources Abstracts from U. S. Geological Survey, and CAIN from the U. S. Department of Agriculture.

continued

Other tape services in areas such as physics, computer science, psychology, and engineering fields will be added as they become available from suppliers.

Current awareness as well as retrospective searches are available on all the sources listed as taped. Interested persons are directed to the contact person.

SAMPLE INPUT

Sample Input is not included for this program. The Sample Output consists of samples of the information returned from a search. Interested persons are directed to the contact person for further information.

SAMPLE OUTPUT

```
PROFILE NO. 2480-008
ACCOUNT NO.                                 ,       THRESHOLD WT. 000000

DATA BASE SEARCHED  CACOND VOLUME 73 ISSUE 18

                                PROFILE

GROUP    TERM    TYPE   WEIGHT              TERM
          NO.

G001      1      TXT     00000     AIR*
G001      2      TXT     00000     EXHAUST*

G002      3      TXT     00000     POLLUT*
G002      4      TXT     00000     CONTAMIN*
G002      5      TXT     00000     WASTE*

G003      6      TXT     00000     SAMPL*
G003      7      TXT     00000     TEST*
G003      8      TXT     00000     MONITOR*
G003      9      TXT     00000     ANALYS*
G003     10      TXT     00000     DETERMIN*
G003     11      TXT     00000     CONTROL*
G003     12      TXT     00000     INSTRUMENT*
G003     13      TXT     00000     COLLECT*
G003     14      TXT     00000     REMOV*
G003     15      TXT     00000     MAK*
G003     16      TXT     00000     MEASUR*
G003     17      TXT               ANAL
G003     18      TXT               DETN

G004     19      CXC               37015

((G001 & G002) | G004) &¬ G003

NUMBER OF ANSWERS TO THIS QUESTION  7



PROFILE NO. 2480-008


GEORGII HW,
UNIVERSITAETSINST.METEOROL. GEOPHYS.,FRANKFURT/M., GER.

****************************************************************

PROFILE NO. 2480-008              THURSDAY, NOVEMBER 5, 1970
                                            QUESTION WT. 0

GEORGII HW,
UNIVERSITAETSINST. METEOROL. GEOPHYS., FRANKFURT/M., GER.

    NATURALLY OCCURRING AEROSOLS IN PURE AND POLLUTED AIR

SCHRIFTENR. VER. WASSER-, BODEN-, LUFTHYG., BERLIN-DAHLEM(SVWLAE)
 1970, NO. 30,() 13-18

REVIEW AEROSOLS POLLUTED ATM

CA-CONDENSATES(CHABA8) 1970, 073(18) 090922

****************************************************************

PROFILE NO. 2480-008              THURSDAY, NOVEMBER 5, 1970
                                            QUESTION WT. 0

CLARENBURG LA,

    AIR POLLUTION IN VLAARDINGEN. CRITIQUE OF THE REPORT OF CLARE
    NBURG                                  ,

CHEM. WEEKBL.(CHWEAP) 1970, 66(31) 16-22

INDUSTRIAL STACKS SULFUR DIOXIDE DOMESTIC HEATERS AIR POLLUTION

CA-CONDENSATES(CHABA8) 1970, 073(18) 090951

****************************************************************
```

COST ESTIMATE

Service is offered to academic institutions only.

Current Awareness Searches

Price per profile per issue

Data Base	*Price*
CA-Condensates	
odd and even	$10.
odd or even only	5.
Chemical Titles	5.
Chemical-Biological Activities	5.
Biological Abstracts	
BA	5.
BIORI	5.
Nuclear Science Abstracts	5.
Compendex	5.

Retrospective Searches*

Price per profile

Data Base	*Price*	
	Volume**	Year
CA-Condensates		
odd and even	$ 70.	$140.
odd or even only	35.	70.
BA-Previews		
BA and BIORI	105.	105.
BA only	70.	70.
BIORI only	35.	35.
CBAC	35.	70.
CT	70.	70.
Compendex	70.	70.
NSA	70.	70.
RIE (through 1970)		75.
CIJE	35.	35.

*Two or three week turn-around schedule

**Volumes correspond to the publishers' printed publication volumes.

continued

Prices are based on profiles with a maximum of 100 search terms. Output on 4x6 inch card stock instead of computer paper is available at an additional cost of 2¢ per card.

Prices are subject to change without notice.

Charge to user = computer costs + network overhead

CONTENTS—INFORMATION RETRIEVAL SERVICES

pages

1– 2	Identification & Abstract
3– 6	User Instructions
7– 8	I/O
9–10	Cost—Contents

227

DESCRIPTIVE TITLE	University Computing Facility
INSTALLATION NAME	Marquette University Computing Center
LANGUAGES	FORTRAN III, FORTRAN IV, COBOL, MAP
COMPUTER	IBM 7040 with 16K-word core storage
AVAILABILITY	Open shop
CONTACT	Mr. Robert Miller, Acting Director, Computing Center, Marquette University, 1515 W. Wisconsin Ave., Milwaukee, Wisc. 53233 Tel.: (414) 224-7700

FUNCTIONAL ABSTRACT

The IBM 7040 at Marquette University is offered to EIN users with Computer Center personnel available for consultation in utilizing the Program Library and in correcting user-written programs.

Equipment at the Marquette University Computing Center includes the following.

Controllers

IBM 1414 Model I	Synchronizer for tapes
IBM 1414 Model IV	Synchronizer for cards, printer

Storage

IBM 729	Tape drives (2)
IBM 7330	Tape drives (4)

I/O

IBM 1402	Card reader-punch
IBM 1403	Line printer

Peripheral equipment (not connected to the 7040)

IBM 407	Tabulating machine
IBM 514	Reproducing punch
IBM 083	Sorter
IBM 870	Graph plotter
IBM 056	Verifier
IBM 1232	Optical reader

continued

Peripheral equipment (continued)

IBM 534	Card punch
IBM 026	Printing card punches(8)
GDDR 3B	Gerber data reader
Microdyne	Analog-to-digital converter

Magnetic tape is available on a loan basis in 200-, 1200-, and 2400-foot sizes. Programs available at Marquette include those developed at the Computing Center and programs ordered through SHARE, with an ordering time of two to three weeks for complete documentation on the latter.

USER INSTRUCTIONS

The Marquette University Computing Center is available on a schedule arranged after individual conference with the EIN Technical Representative. Charges for computational services for small colleges may be waived. A potential user should first fill an Account Initiation Form with the EIN Project Office. EIN will subsequently notify the Technical Representative at Marquette, who will contact the potential user and arrange for a project code number.

To aid the Computing Center user in programming and usage, the publications listed in Ref. 1–8 are of general assistance. The program library[9] should be consulted for specific program information. In addition, the Computing Center maintains a limited reference library consisting primarily of periodicals published by computer manufacturers, professional organizations and other educational institutions.

The Computing Center personnel are available for consultation, programming assistance, and keypunching on a limited basis without charge; extensive utilization of these services can be obtained on a contracted basis.

REFERENCES

1. *Users Manual, MUCC #00*, (Marquette Univ., Milwaukee, Wisc., Sept., 1968).

2. *Programmers Guide, MUCC #10*, (Marquette Univ., Milwaukee, Wisc., Sept., 1967).

3. *Operators Guide, MUCC #12*, (Marquette Univ., Milwaukee, Wisc.).

4. *Autoplotter Programmers Manual (Peter's Version), MUCC #110*, (Marquette Univ., Milwaukee, Wisc.).

5. *FORTRAN III Reference Manual, MUCC #111*, (Marquette Univ., Milwaukee, Wisc.).

6. *Survey Report Generator (SRG), MUCC #112*, (Marquette Univ., Milwaukee, Wisc.).

7. *FORTRAN IV Error Messages, MUCC #200*, (Marquette Univ., Milwaukee, Wisc.).

continued

8. *Test Score and Analysis System, MQ070*, (Marquette Univ., Milwaukee, Wisc.).

9. Copies of all programs and documentations can be obtained by inquiring to the EIN Technical Representative at Marquette University. An index can be obtained either from the Technical Representative or through the EIN Project Office at the cost of reproduction and mailing.

PRICING

Description	*Rate per hour*
Computation	$240.00
Consultation	$ 10.00
Programming	$ 15.00
Keypunching	$ 2.50

Note: Computation charges may be waived for small colleges usage on a non-regular basis. Interested persons should contact the EIN Technical Representative.

CONTENTS

pages
1– 2 Identification & Abstract
3– 4 User Instructions
5 Pricing—Contents

DESCRIPTIVE TITLE	Kiewit Computation Facility
INSTALLATION NAME	Kiewit Computation Center Dartmouth College
LANGUAGES	Dartmouth BASIC Dartmouth FORTRAN Dartmouth ALGOL Dartmouth LISP LAFFF (Language to Aid Financial Fact Finders) CRIII (Computer Research Involving Investment Information) MIX GEFORT (GE FORTRAN) GMAP (GE Macro Assembly Program) TRAC (Text Reckoning and Compiling Program)
COMPUTER	GE-635
AVAILABILITY	Available to specified categories of remote terminals
CONTACT	For Account Initiation, Period of Usage, and Policy: A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

For specific problems in the following areas (all extensions preceded by Area Code 603, exchange 646-):

Area	*Name*	*Extension*
Validating user numbers and passwords; granting special permissions	Nancy Broadhead Stephen V.F. Waite	2643 2643
Programming assistance	Diane Mather Stephen V.F. Waite	3283 2643
Use of Plotters	Arthur Luehrmann	2976,2864
Supplying Manuals	Jann Dalton	2147

continued

FUNCTIONAL ABSTRACT

The Dartmouth Time Sharing System is a remote access facility supporting such low-speed terminals as Teletype models 33, 35, and 37, Friden 7100 and IBM 2741. Detailed terminal, coupler, and bit rate specifications are contained in references 1 and 2. On-site peripherals are not available to remote users, since the Center does not handle input or output beyond the operations counter.

Hardware available to remote users includes the dual processor GE 635 with 160K x 36-bit words; two GE Datanet 30's with a total of 155 bit buffer units; one GE magnetic drum with controllers; two IBM 2314 disc units with a Datametrics interface to a GE IOC controller; and six tape drives (7 track) operated by two tape controllers.

The executive system was written and is maintained and modified by Dartmouth undergraduates under the supervision of the Software Development Director at Kiewit.

Remote faculty users will be alloted 16K of core memory, 64 seconds of run-time, and, in most cases, 6K of disc storage.

Special software available to all users includes the LAFFF and the CRIII systems developed by the Amos Tuck School of Business to aid in finding financial facts (see Ref. 3); a simulator of the MIX system designed by Donald Knuth; RUNOFF (for text processing); Dartmouth EDIT, TEXTEDIT, and STRING EDITOR, all for file manipulation, whether the file be a program, text, or alphanumeric data; TEACH, which provides the possibility for writing a program to automatically test programs written by students for a given problem or course; and IMPRESS, which consists of numerous sociological data files and a core of programs for manipulating them.

REFERENCES

1. Hargraves, R., and Mather, D., *User's Guide to the Dartmouth Time-Sharing System*, Dartmouth College TM 022 (Hanover, N.H.: Oct. 1970).

2. Morton, A.K., *Hardware and Access Supplement for Remote Users* (Dart. Col., Jan. 1971).

3. *CRIII* (Tuck Sch. of Bus. Admin., Dart. Col., 1970). CRIII is a system for financial fact finding written in BASIC. It is now being used almost exclusively with and by Tuck School students, and is obtainable from the Tuck School.

USER INSTRUCTIONS

Procedure

The user should first file an Account Initiation Form with the EIN Office, after which he will be issued a valid user number and password to be recognized by the validation monitor from the starting date to the ending date (inclusive) cited under "Account Limits" on the Account Initiation Form. Should the user desire to extend this period, he should contact the Technical Representative at Dartmouth, either personally or through his own Technical Representative. Should the TR at Dartmouth not be immediately available, the user may contact one of the persons responsible for validation and permissions, as indicated under the section CONTACT.

Availability

The Dartmouth Time Sharing System is available to all users according to the following schedule.

Monday	8:00 a.m. – 1:00 a.m.
Tuesday	8:00 a.m. – 1:00 a.m.
Wednesday	8:00 a.m. – 8:00 p.m.
Thursday	8:00 a.m. – 1:00 a.m.
Friday	8:00 a.m. – 1:00 a.m.
Saturday	8:00 a.m. – 1:00 a.m.
Sunday	8:00 p.m. – 1:00 a.m.

Services

The staff of the Computation Center does not provide programming services for outside users, except by way of consultation in specific problem areas. Consulting on specific problems is provided on a limited basis without surcharge as a routine user service.

Use of on-site peripherals is not permitted to remote users as the Center does not handle physical input and output through the mail. Should remote use of a plotter be required, DTSS software will presently drive CalComp Model 565 and Timeshare Devices, Inc. Model C/P 701.

Informational Services

The following files are maintained in the Dartmouth library to keep users informed of current procedures and scheduling, new system features, and revisions:

continued

DTSSINFO	files under this heading contain general information of interest to DTSS users
CCNEWS	Computation Center and system news
PLOTNEWS	information pertaining to the use of the library routines for the Timeshare Devices plotter
PINEWS	information of interest to users of Project IMPRESS

Information pertaining to the use of the library is contained in Ref. 1. Log-on procedures and other information are described in Refs. 1 and 2.

REFERENCES

General

1. Hargraves, R., and Mather D., *User's Guide to the Dartmouth Time-Sharing System*, Dartmouth College TM 022 (Hanover, N.H.: Oct. 1970).

2. Morton, A.K., *Hardware and Access Supplement for Remote Users* (Dart. Col., Jan. 1971).

3. *BASIC*, (Kiewit Comp. Center, Dart. Col., Sept. 1970), 5th ed.

4. Mather, D., and Hart, R., *A User's Guide to the DTSS Program Library*, Dart. Col. TM 010 (Oct. 1970).

5. Dobbs, G., *Dartmouth EDIT*, Dart. Col. Tm 002 (Jan. 1969).

6. Takaro, T., *Dartmouth STRING EDITOR*, Dart. Col. TM 003 (Sept. 1969).

7. Takaro, T., *TEXT EDITOR User Manual*, Dart. Col. TM 004 (Sept. 1969).

8. Pataky, M. and Hills, D., *RUNOFF User's Reference Manual*, Dart. Col. TM 005 (Jan. 1970).

9. Hills, D., *XRUNOFF Manual: A Supplement to the RUNOFF User's Reference Manual* (Dart. Col., 1970).

10. Hargraves, R., *Double Precision in Dartmouth BASIC*, Dart. Col. TM 006 (July 1969).

11. Hargraves, R., and Semprebon, L., *CalComp Programming in BASIC*, Dart. Col. TM 007 (Apr. 1969).

continued

12. Garland, S., and McGeachie, J., *Multiple Terminal Programming*, Dart. Col. TM 009 (Nov. 1969).

13. Marshall, S., *An ALGOL 68 Garbage Collector*, Dart. Col. TM 011 (Dec. 1969).

14. Reiss, S., and Behrens, A., *MIX User's Reference Manual*, Dart. Col. TM 013 (Dec. 1969).

15. Luehrmann, A., *Use of the Timeshare Devices Plotter on DTSS*, Dart. Col. TM 014 (Jan. 1970).

16. *Dartmouth Time-Sharing Fortran* (1971); adapted from the General Electric Inform. Serv. Depart. Ref. Manual *FORTRAN Language*, Document No. 802209C, and printed with the permission of G.E.

17. Kurtz, T., *Accurate Matrix Inversion in BASIC*, Dart. Col. TM 016 (Apr. 1970).

18. Hobbs, S., *LISP System Reference Manual for DTSS on the GE 635*, Dart. Col. TM 017 (July 1970).

19. McGeachie, J., *Use of the IBM 2741 Communications Terminal on DTSS*, Dart. Col. TM 018 (Aug. 1970).

20. Relson, D., *Dartmouth ALGOL for DTSS*, Dart. Col. TM 021 (June 1970).

21. Kemeny, J.G., and Nevison, J.M., *How to Write a TEACH Program* (Dart. Col., Dec. 1967).

22. Noma, E., *Communications Facilities Used by Kiewit*, Dart. Col. TM 026 (Dec. 1970).

Amos Tuck School of Business Administration

23. *CRIII* (Tuck Sch. of Bus. Admin., Dart. Col., 1970).

Project IMPRESS

24. Meyers, Jr., E.D., *An Introduction to Project IMPRESS*.

25. *IMPRESS Files: Structure and Format* (Dec. 1969). This paper describes the format of IMPRESS data files in the computer.

continued

26. *PRES 68* (Oct. 1969, first print.; and Jan. 1970, second print.). Codebook for the 1968 Presidential Election Survey, Univ. of Mich. Survey Research Center Study, SRC S523.

27. *DART 69* (Apr. 1970). Codebook for the 1969 Dartmouth Coeducation Survey.

28. *WORK* (Apr. 1970). Codebook for *Organization of Work*, a comparative analysis of production among nonindustrial peoples, Stanley H. Udy, Jr., (New Haven: HRAF Press, 1959).

29. *DART 70* (May 1970). Codebook for the 1970 Dartmouth Coeducational Survey.

Notes

1. All technical memoranda (TM), as well as the remainder of references 1-22, are available from the Documents Center at Kiewit Computation Center, Dartmouth College. Users may also request to be placed on the DTSS Phase II Documentation List, whereby they will receive, for an annual fee, all system documentation as it appears, including revisions. This documentation includes information on GMAP and the TRAC, LISP, FORTRAN, ALGOL, BASIC, and EDIT systems.

2. CRIII and LAFFF are parallel systems for financial fact finding, the former being written in BASIC and the latter in machine language. LAFFF Manuals are no longer available; so information on the use of this system must be obtained by writing the Tuck School at Dartmouth College. CRIII is now being used almost exclusively with and by Tuck School students, and is obtainable from the Tuck School.

3. There is no system manual for TRAC. The implementation, however, is fairly standard, and may be understood by consulting almost any reference on the language.

4. Ref. 24-29 are available from IMPRESS Office, Silsby Hall, Dartmouth College.

COST ESTIMATE

	Educational	*Commercial*
Terminal Use		
110 BAUD	3.50/hr.	11.00/hr.
134.5 BAUD	4.35/hr.	12.75/hr.
150 BAUD	4.35 hr.	12.75/hr.
300 BAUD	(not yet available)	
Central Processor Use	.11 sec.	.35 sec.
Storage (1000 word units)	2.95/unit	3.00/unit

Charge to user = computer costs + communication + EIN network overhead

CONTENTS—KIEWIT COMPUTATION FACILITY

pages	
1–2	Identification & Abstract
3–6	User Instructions
7	Cost—Contents

DESCRIPTIVE TITLE	Remote Job Entry System (RJE)
INSTALLATION NAME	The Pennsylvania State University Computation Center
LANGUAGE	Full facilities of the IBM Operating System for the 360.
COMPUTER	IBM 360/67
AVAILABILITY	Remote access
CONTACT	Dr. Daniel Bernitt, The Pennsylvania State University Computation Center, 105 Computer Building, The Pennsylvania State University, University Park, Pa. 16802 Tel.: (814) 865-9527

FUNCTIONAL ABSTRACT

The Remote Job Entry (RJE) system permits research users to access the System 360/67 computer from a remote keyboard terminal in a time-sharing mode. The terminals supported by the system are IBM 2741's with BCD or "SELECTRIC" character set, DATEL or other terminals that simulate the 2741 and hence have the same characteristics, IBM 1050 terminals with the BCD character set, and Teletypes.

Principal Equipment

Main Campus

Number	*Type*	*Use*
1	IBM 360/67	Main system
5	IBM 360/20	High speed entry
15	IBM 1050	Low speed entry
18	IBM 2741	
20	DATEL Thirty-21	
1	IBM 2250	Interactive Graphics
1	CalComp Plotter (30 inch bed)	Hard Copy Graphics

Branch Campuses

Number	*Type*	*Use*
10	IBM 2780	High speed entry
1	IBM 360/20	High speed entry
3	IBM 1050	Low speed entry

continued

Storage and I/O Devices

The 360/70 central system consists of: 1 million bytes main memory; 2 million bytes LCS; six IBM 2400 tape drives (9-track); two IBM 2400-1 tape drives (7-track); two IBM 2301 Drums; three IBM 2314 Disk Drives: one IBM 2703 Communications Adapter; two IBM 2540 Card reader/punches; three IBM 2403 Printers; one IBM 2671 Paper Tape reader; five IBM 2260 CRT.

Software

1. Complete OS/360 (currently Release 18 + HASP).
2. Locally developed remote job entry, supporting low speed terminals of the type

   IBM 2741 (EBCDIC or SELECTRIC character set) or
   DATEL THIRTY-21 (portable, similar to 2741's) or
   IBM 1050 (EBCDIC character set).
3. Various language processors and application programs. Includes systems developed locally and at other independent computing facilities along with all IBM type I and II systems.

REFERENCES

The Pennsylvania State University Computation Center, *Remote Job Entry System: General Information and Summary Description* (University Park, Pa.: The Penn. State Univ., Nov. 1970). Available from the EIN Office at the cost of reproduction and mailing.

USER INSTRUCTIONS

Interested persons should consult the contact person and the reference listed below.

REFERENCES

The Pennsylvania State University Computation Center, *Remote Job Entry System: General Information and Summary Description* (University Park, Pa.: The Penn. State Univ., Nov. 1970). Available from the EIN Office at the cost of reproduction and mailing.

COST ESTIMATE

Charge for computer time = $0.11/sec.

In addition, when the RJE system is used, there are two types of specific charges. First is the charge for usage of a typewriter terminal, based on actual connection time (time between "logging on" and "logging off"). The time is calculated to the nearest minute and charged at $0.07 min.

The second charge is for the use of storage space, computed on the basis of the number of files (sets of 500 eighty-column card images) belonging to a user and the number of days each file is retained. The retention of files is charged at the rate of $0.12/file/day.

Charge to user = computer costs + network overhead

CONTENTS—REMOTE JOB ENTRY SYSTEM (RJE)

pages	
1– 2	Identification & Abstract
3	User Instructions
5	Cost—Contents

DESCRIPTIVE TITLE	Vogelback Computing Facility
INSTALLATION NAME	Vogelback Computing Center Northwestern University
LANGUAGES	FTN (CDC FORTRAN Extended) RUN (CDC FORTRAN) COBOL, SNOBOL, LISP, ALGOL, MIMIC, SLIP, SIMSCRIPT, COMPASS
COMPUTER	CDC 6400, 65K
AVAILABILITY	Batch processing, remote batch processing, remote on-line processing
CONTACT	Lorraine Borman, EIN Technical Representative, Vogelback Computing Center, Northwestern University, 2129 Sheridan Road, Evanston, Ill. 60201 Tel.: (312) 492-3682

FUNCTIONAL ABSTRACT

The 6400 is primarily a batch processing machine but can handle on-line access through Teletype models 33 or 35, or a CDC 200 User Terminal.

Equipment

1	6416 Central computer (65K memory)
1	6603 disk
2	854 disk pack drives
2	501 printers (upper case only)
1	415 card punch
2	405 card readers
1	565 CalComp plotter
4	607 magnetic tape transports (7 track)
1	3691 paper tape reader/punch
1	6671 communications multiplexor

The Center is currently running under the SCOPE 3.2 operating system. Data tapes must be 7 track, BCD; only 026 character set is acceptable on either punched cards or tapes.

Use of the Center is restricted to educational and non-profit institutions.

USER INSTRUCTIONS

The Center is open for job processing from 9:30 A.M. to midnight, Monday through Friday, and 9:30 A.M. to 5:00 P.M. on Saturdays. Consulting services will not generally be available for EIN users.

REFERENCES

CDC 6400 User's Manual (Evanston, Ill.: Vogelback Comp. Center, Northwestern Univ., Feb. 1968). Available from Vogelback Comp. Ctr. at a cost of $3.25.

ONLINE Writeup, VCC Doc. No. 195 (Evanston, Ill.: VCC, Northwestern Univ.). Available without charge.

CDC System and Language Manuals. Individually priced; list available upon request.

COST ESTIMATE

Charges are based on the amount of central processor and peripheral processor time used by each job, the amount of central memory occupied, and for on-line jobs, a connect charge of $1.00/hour (maximum charge, $40/month).

Charging for central processor (CP) and peripheral processor (PP) time used on all ordinary jobs is at a rate of $7.50, $8.50, or $9.50 per minute, depending upon the central memory field length (FL) partitions each job utilizes during processing. Chargeable minutes are calculated as the sum of central processor plus 20% of peripheral processor time. In detail, the charging algorithm is as follows:

$$\text{Total Job Charges} = \sum_{i=1}^{3} R_i \, (CP_i + 0.2\ PP_i)$$

where R_1 = \$7.50 per minute for $0 < FL \le 43{,}000_8$ $(17{,}920_{10})$

R_2 = \$8.50 per minute for $43{,}000_8 < FL \le 100{,}000_8$ $(32{,}768_{10})$

R_3 = \$9.50 per minute for $100{,}000_8 < FL \le 140{,}000_8$ $(49{,}152_{10})$

and where CP_i and PP_i are the amounts of central processor and peripheral processor time recorded in each central memory partition during job processing.

A long job is defined as one which:

1. requires over $70{,}000_8$ field length and over 10 minutes of CP + PP time

Or 2. requires over 200 pages printed

Or 3. requires over 2000 cards punched.

Long jobs are generally run only after midnight. The long-job rates are $6.50, $7.00, and $7.50 per minute for the three central memory partitions. Jobs which are in any of the above three categories *must* be run as long jobs or as P 0 jobs (see below). If a long job is submitted for normal processing, job execution or output will be terminated. Users who wish to take advantage of long job charging may submit a normal job as a long job.

continued

A zero-priority job (P 0) is one that will be processed only after all other jobs have been completed (including long jobs). P 0 jobs will have a lower priority than Computing Center systems projects and, thus, will not usually be run after 3:00 A.M. Therefore, 24-hour turn-around time will not be guaranteed. The zero-priority rates are \$5.00, \$5.50, and \$6.00 per minute for the three central memory partitions.

The minimum job charge is \$1.00 for jobs submitted at the central site and \$0.50 for jobs submitted through remote terminals.

EIN users will be charged an additional \$15.00/job processed to cover local overhead.

Charge to user = computer time + postage and handling + network overhead
= computer time + \$15.00 + network overhead

CONTENTS—VOGELBACK COMPUTING FACILITY

pages

1 Identification & Abstract
3 User Instructions
5– 6 Cost—Contents

DESCRIPTIVE TITLE	MERC Time-Sharing System
INSTALLATION NAME	Middle-Atlantic Educational and Research Center
LANGUAGES	FORTRAN IV, COBOL, ASSEMBLY, BASIC, and various special applications programs.
COMPUTER	RCA Spectra 70/46G with 256 kilobytes of core
AVAILABILITY	Remote and mail access
CONTACT	Dr. Richard S. Lehman, Director MERC, P. O. Box 1372, Lancaster, Pa. 17604 Tel.: (717) 393-0132 or 393-5021 Dr. Paul W. Ross, Manager of Services, MERC, P. O. Box 1372, Lancaster, Pa. 17604 Tel.: (717) 393-0132 or 393-5021

FUNCTIONAL ABSTRACT

MERC is making available its RCA Spectra 70/46G with 2-megabyte drum, four disc storage units and four tape drives (9-channel, 800 bpi). This system is a highly interactive time-sharing and simultaneous batch computer operating system. The terminals supported are Model 33 and 35 teletypewriters, via 1/2 state WATS lines in the 215 and 717 dialing areas. Local dial-up in the Lancaster area and other selected locations in eastern Pennsylvania and southern New Jersey is available.

The system supports programs in BASIC, FORTRAN IV, COBOL, and Assembly. The FORTRAN system has both background compilation and execution, fast batch and interactive FORTRAN. COBOL is supported for conventional background batch-mode operation or in a mode in which input may be taken from a terminal. In addition, a syntax checking system is available.

Entry of data is via paper tape reader at the user's site, or by input of card decks at the computer center in Lancaster, for building large data base files. In addition, the system provides a large data base management system known as QWIK-TREIVE for special on-line data management applications. Various other special application packages are available.

continued

REFERENCES

Coia, E.J., *Interactive BASIC System Reference Manual* (Lancaster, Pa.: MERC, July 1970).

MERC Program Library Manual (Lancaster, Pa.: MERC, July 1970).

Gallagher, J., *MERC Calculator Manual* (Lancaster, Pa.: MERC, July 1970).

Copies of these manuals are available from MERC or through EIN at the cost of reproduction and mailing.

USER INSTRUCTIONS

Instruction is provided in the form of various manuals. The system for BASIC use is represented by Ref. 1. FORTRAN IV is the conventional version with very minor modifications to meet the requirements of our system. Output may be directed to the user terminal or, under certain circumstances, may be directed to the printer located at the MERC site. The special applications are described in the MERC library manual.[2]

The system is available at the following times.

8:00 a.m. to midnight, Monday through Friday
10:00 a.m. to 6:00 p.m., Saturday
Other hours by special arrangement
Normal preventive maintenance and system unavailable, 8:00 a.m. to 10:00 a.m., Tuesday and Friday.

REFERENCES

1. Coia, E.J., *Interactive BASIC System Reference Manual* (Lancaster, Pa.: MERC, July 1970).

2. *MERC Program Library Manual* (Lancaster, Pa.: MERC, July 1970).

3. Gallagher, J., *MERC Calculator Manual* (Lancaster, Pa.: MERC, July 1970).

Copies of these manuals are available from MERC or through EIN at the cost of reproduction and mailing.

COST ESTIMATE

CPU Time	05¢ per CRU*

*A CRU is defined as CPU seconds times a program weighting factor, which is dependent upon the characteristics of the program.

Connect Time

110 baud to 150 baud	$ 9.00/hr.
300 baud	$12.00/hr.
1200 baud	$15.00/hr.
2000 baud	$20.00/hr.

Peripheral Time

Card reading	$0.20/1000 cards
Card punching	$2.00/1000 cards
Line printing	$0.75/1000 lines
Private volume usage (tape & disk)	$0.10/minute

Storage Charges

On Line Storage (590 Disks)

First 25 pages*	No charge
Next 75 pages	$1.00/page
Next 400 pages	$0.70/page
Next 500 pages	$0.50/page
Above 1000 pages	$0.25/page

*A page is 2048 characters (8 bit + parity) of storage

Off Line Storage (Magnetic tape)

MERC supplied tapes	$2.00/month
Customer supplied tape	$1.00/month

Guaranteed Minimum Discount Policy

If the user is willing to guarantee MERC a monthly minimum billing, MERC will give the user a discount on computer charges (i.e. CPU TIME, CONNECT TIME, AND PERIPHERAL TIME) in the amount of (minimum guaranteed by customer/100)% up to a maximum of 30%.

continued

Educational users fall under the following minimum discount policy. There is no maximum to the amount of discount available and the discount is calculated as:

$$(\sqrt[3]{\text{minimum guaranteed/month}})\%$$

Educational Discount

Because of MERC's dedication to the educational community, a very liberal discount of 50% is allowed to educational institutions. This discount is applied to the total billing after any guaranteed minimum discounts are applied. However, educational institutions electing the guaranteed minimum discount will be expected to make the minimum amount after all discounts are applied.

Costs of manuals, consulting, supplies, and other support costs are not covered by the above discount policies.

Charge to user = computer costs + network overhead

CONTENTS—MERC TIME-SHARING SYSTEM

pages

1– 2 Identification & Abstract
3 User Instructions
5– 6 Cost—Contents

DESCRIPTIVE TITLE	UCSB On-Line System
CALLING NAME	OLS
INSTALLATION NAME	University of California Santa Barbara UCSB Computer Center
AUTHOR(S) AND AFFILIATION(S)	Dr. David Harris, Principal Investigator Ron Stoughton, Principal Programmer UCSB Computer Research Lab.
LANGUAGE	Basic Assembler Language
COMPUTER	IBM System 360/75
PROGRAM AVAILABILITY	Listing of programs may be obtained by special arrangement with the Computer Research Lab.
CONTACT	Glenn Davis, OLS Manager, Computer Center, Univ. of Cal., Santa Barbara, Cal. 93106 Tel.: (805) 961-2462

FUNCTIONAL ABSTRACT

The UCSB On-Line System (OLS) provides the capability for sophisticated mathematical analysis for use in solving problems where human interaction is either necessary or desired.

OLS accepts both real and complex numbers (scalars) as operands as well as lists of such numbers (vectors). Operations performed on scalars produce scalar results, which can be numerically displayed; operations on vectors produce vector results (the specified operation being performed on each component), and results of computation can be displayed either numerically or graphically. Operands can be stored and used as required. Operators include sine, cosine, logarithm, and exponentiation; and each is executed with a single button push. Facility is provided for interaction between operands of different types (e.g. vectors and scalars). In addition, a limited set of operations manipulate integers used in subscripting.

Additional features are provided to support OLS's basic mathematical capability. Although OLS normally executes each

continued

button as it is pushed, a button sequence may be defined, named, and saved for later execution. Convenient means are provided for editing such sequences. Lists of buttons to be executed can include programmed pauses, allowing manual and programmed activity to be interfaced; as well as branching based upon results of computation. Messages can be composed of alphameric, Greek, and special characters, and displayed. Those characters not specifically provided by OLS may be designed by the user and stored, and then are available for use. A collection of button lists and user-created characters is referred to as a "system". Systems are named and can be permanently stored and later retrieved. Portions of systems may be transferred between systems, and systems may be transferred between users. Sets of scalars and vectors may also be named, permanently stored, and later retrieved.

Apart from OLS's mathematical capability, a recent development provides the ability to create and edit a "deck" of cards and submit it for execution in an OS partition. Operations on string, record and file levels are provided. Data-sets residing on any disk pack within the installation may be fetched, examined, modified, and submitted for execution. Work continues in the general field and further developments are expected.

USER INSTRUCTIONS

To use OLS, a terminal consisting of the following components is required:

MESG keyboard
Storage display tube
Telephone interface & display controller
2400 bps lines, 201B data sets, etc.
Digital plotter (optional)
Teletype printer (optional)
Grafacon-graphic input unit (optional)

OLS is available to users 0900 hours to 2200 hours daily and 1200 hours to 1800 hours on Saturday. The remaining hours are designated as development time but the system will be available if no development is in progress.

A consultant is available to assist users weekdays 0900 hours to 1700 hours.

REFERENCES

UCSB On-Line System Manual (Santa Barbara: UCSB Computer Center, Jan. 1, 1971).

Howard, J.A., *Mini-Users-Manual for UCSB On-Line System* (UCSB, July 1969). Adapted from script of On-Line System video tape.

Bruch, J.C., Jr., "Free Streamline Theory and Computer Generated Displays," artical submitted for review and publication in the ASCE *J. of Eng. Mech. Div.* (Aug. 1970).

Bruch, J.C., Jr., *Hydrodynamics Through On-Line Computer Generated Displays* (Handbook), (UCSB Col. of Eng., xeroxed).

Bruch, J.C., Jr., "Two Dimensional Flow Visualization Using Computer Generated Displays," article submitted for review and publication in the ASCE *J. of the Hydraul. Div.* (Aug. 1970). Concerning 16mm educational/demonstration movie, Learning Resources Motion Picture Prod. Sec. UCSB, 1970.

Bruch, J.C., Jr. and Howard, J.A., "Flow Visualization in Hydrodynamics Using On-Line Generated Displays, paper presented at the Pacific Southwest ASEE Annual Meeting, Univ. of Cal., Davis, Dec. 29-30, 1969. Concerns 16mm movie, Learning Resources Motion Picture Prod. Sec., UCSB, 1969.

continued

Bruch, J.C., Jr. and Wood, R.C., "The Teaching of Hydrodynamics Using Computer Generated Displays," Bull. Mech. Eng. Ed., *9* (May 1970), pp. 105-115.

Economics Data Bank Handbook (UCSB Econ. Lab., Nov. 1969).

Economic Statistics: Handbook for the On-Line System (UCSB Econ. Lab., Dec. 1969).

Macro-and-Micro Economic Models (Handbook), (UCSB Econ. Lab., Oct. 1969).

Purpose and Functions (Report), (UCSB Econ. Lab., Sept. 1969).

Sample Instructional Materials (UCSB Econ. Lab., Nov. 1969).

Teleputer Handbook (UCSB Econ. Lab., Dec. 1969), Revised.

Ewig, C.S., Gerig, J.T. and Harris, D.O., *An Interactive On-Line Computing System and its Use in Chemistry Education* (UCSB Depart. of Chem., xeroxed).

Howard, J.A. and Wood, R.C., "Computer-Assited Instruction in Engineering Using On-Line Computation," *J. of Eng. Ed.*, (in press).

Kinsey, P., Ewig, C.S. and Harris, D.O., *Introduction to the UCSB On-Line Computing System* (UCSB Depart. of Chem., 1970).

Sandusky, A., *Handbook for the Computer Laboratory: Analysis of Data in Psychology* (UCSB Depart. of Psych., July 1, 1970).

Sandusky, A., *Undergraduate Researcher's Manual: Analysis of Data in Psychology* (UCSB Depart. of Psych.).

Sullivan, J.J., "Computer Based Instruction in Economics: A Report on Facilities and Applications at UCSB," paper presented at a Conference on Computers in Undergraduate Curricula, Univ. of Iowa, Iowa City, June 16-18, 1970.

Sandusky, A., "The Economics Laboratory at UCSB," Simulation and Games: an Intern. J. of Theory, Design, and Res., *1*, 1 (March 1970), pp. 81-91.

Wood, R.C. and Bruch, J.C., Jr., "Teaching Complex Variables with an Interactive Computer System," article submitted for review and publication in the *IEEE Trans. on Ed.* (July 1970).

Wood, R.C., and Howard, J.A., "An Interactive Computer Classroom," Ed. Res. and Methods J., *2*, 4 (June 1970), pp. 29-31.

COST ESTIMATE

Charges for use of the On-Line System are the sum of the following.

Description	*Price*
Connect time	$4.00
Central processor unit	0.0001793/CRU[α]
Core units	0.1776/K-byte-hr
Disk storage	0.10/K-byte-month

[α] A CRU is defined as CPU seconds X a program weighting factor which is dependent on the characteristics of the machine.

The cost per hour varies from user to user but the average for all users is $8.40. Student usage is nearly always on the low side of the mean cost/hr.

Charge to user = computer costs + EIN network overhead

CONTENTS—OLS

pages	
1– 2	Identification & Abstract
3– 4	User Instructions
5	Cost—Contents

DESCRIPTIVE TITLE	Individual Case Statistics Program
CALLING NAME	SUMSCRDS
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine University of Iowa Computer Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

SUMSCRDS is used to calculate a mean, sum, standard deviation, and the number of non-missing observations in one individual case. It handles each case separately. The case may continue for as many cards as necessary (no. variables = 80). Missing data are represented as a blank or minus zero.

USER INSTRUCTIONS

Input Deck

The input deck consists of the following cards in the order described.

Title Card

Columns	*Contents*
1–72	Any alphanumeric information.

Problem Card

Columns	*Contents*
1– 2	Number of variables ($\leq$ 80).
3– 7	Number of cases ($\leq$ 99,999).
8	Number of Variable Format Cards.
9	5: card input. Tape number for special tape input.

Variable Format Card

A maximum of five cards is allowed. Use the standard F-type format statement.

Observation Cards

Order the cards so that one complete case is followed by the next complete case, according to the format statement. A maximum of 99,999 cases is allowed.

Stop Card

Columns	*Contents*
1– 4	STOP

Machine Requirements

The SUMSCRDS program itself is less than 3K bytes long. It uses a special OS Assembler Language program, BLNK.

SAMPLE INPUT

```
TESTING SUMSCRDS PROGRAM
 4    215
(4F2.0)
17121118
10241015
STOP
/*
//
```

SAMPLE OUTPUT

```
 TESTING SUMSCROS PROGRAM
     NO. OF CASES =     2
     NO. OF VARIABLES = 4
     FORMAT STATEMENT(4F2.0)




 TESTING SUMSCROS PROGRAM
     CARD NO.      SUM                 MEAN              STANDARD DEVIATION    NO. OF CASES
                                                                              NOT MISSING
         1     0.580000E 02       0.145000E 02          0.351188E 01           4.
         2     0.590000E 02       0.147500E 02          0.660177E 01           4.
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $0.75.

Charge to user = computer costs + postage + network overhead
= $0.75 + postage + network overhead

CONTENTS—SUMSCRDS

pages
1 Identification & Abstract
3 User Instructions
5– 6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	General Multiple Regression Analysis
CALLING NAME	REGAN1
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Stan Walljasper, IBM 7044 version Converted to IBM 360 by Louise R. Levine The University of Iowa Computer Center
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

REGAN1 computes the Pearson Product Moment correlation matrix, mean, standard deviation, and population standard deviation. If desired, the multiple R, standard error of the estimates, F ratio, degrees of freedom, regression coefficients, partial coefficients and F ratio associated with each individual variate, the intercept constant, and residuals also may be calculated. There is *no* allowance for missing data. The basic equations and a simple explanation may be found in Cooley and Lohnes.[1]

REFERENCES

1. Cooley, W.W., and Lohnes, P.R., *Multivariate Procedures for the Behavioral Sciences*, (John Wiley & Sons, Inc., New York, 1962), pp. 31–45.

USER INSTRUCTIONS

Input Deck

The input deck consists of the following cards in the order described below.

Title Card

This card gives any alphanumeric information to be printed as a heading on the output.

Problem Card

Columns	*Contents*
4– 5	Number of variables (independent + dependent $\leq$ 50)
6–10	Total number of observations
11–15	1: No data error check is desired 2: Data error check is desired
16–20	1: No residuals are to be calculated 2: Residuals are to be calculated
21–30	If data error check is requested, punch the lower bound for allowable data. The assumed format is F10.5. Leave blank if no data error check is requested.
31–40	If data error check is requested, punch the upper bound for allowable data. The assumed format is F10.5.
41	Number of Format Cards ($\leq$ 3)
42	blank: card input 2,3,4,8, or 9: unit number of disc or tape input Appropriate DD cards for this file should be placed before the //GO.SYSIN card.

Format Cards

Columns	*Contents*
1–80	The data should be read by subject for all variables with an E or F type FORTRAN format. A maximum of 3 Format Cards is allowed.

Data Cards

The Data Cards contain the records of observations. Any uniform

continued

number of cards per observations may be used. Each observation must start on a new card.

/* Card

//GO.FT05F00n DD * Card

This card must be present for each problem, with n taking the value of 2 for the first problem and increasing by one for each additional problem.

Permutation Cards

Columns	*Contents*
5	1: No residuals are desired 2: Residuals are desired
6–10	Number of variables to be used for the multiple R (independent + dependent variables)
11–12	Dependent variable number
13–14	Independent variables by number. If there are more than 34 independent variables, continue on another card beginning in Cols. 1–2.
⋮	⋮
79–80	

Only one problem can be run with as many Permutation Cards (max. 20) as necessary.

Program Deck

```
//name JOB (project no., etc.)
//step EXEC STAT,PARM.GO='REGAN1'
//GO.SYSIN DD *
     Title Card
     Problem Card
     Format Card
     Data Cards
/*
//GO.FT05F002  DD *
     Permutation Cards
/*
```

Method

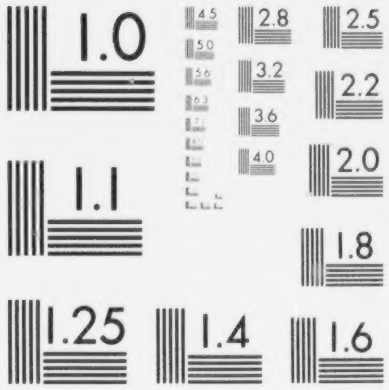
The data are read and the summation on matrices are updated using an auxiliary index matrix.[1] Means, deviations, and product-moment

continued

5 OF 12

ED

064 958


1.0
4.5
5.0
5.6
6.3
2.8
3.2
3.6
4.0
2.5
2.2
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

correlations are computed by a triangular method.[2] The simple correlation matrix is transferred to a work area from which the multiple R, Beta weights, etc., are calculated. The basic equations and a simple explanation may be found in Cooley and Lohnes.[3]

REGAN1 runs on a 360/65 under HASP and OS with FORTRAN IV (G) and uses 70K bytes of core during execution with the source deck. A work area for data set reference number 1 is required.

REFERENCES

1. Caffrey, J. "CRAM, Algorithm 67", *Collected Algorithms from CACM,* (A.C.M., New York).

2. Hafley, W.L., and Lewis, J.S., "Triangular Regression, Algorithm 142", *Collected Algorithms from CACM*, (New York, A.C.M.).

3. Cooley, W.W., and Lohnes, P.R., *Multivariate Procedures for the Behavioral Sciences*, (John Wiley & Sons, Inc., New York, 1962).

SAMPLE INPUT

```
TESTING REGAN1
   4  21    1    2                                     1
(4F10.0)
      68.0        2.53        30.6        3.8
      77.8        2.61        34.0        3.8
      83.8        2.82        46.0        3.8
      83.0        2.66        35.7        5.9
      83.8        2.62        54.1        5.3
      90.5        2.86        59.3        3.3
      92.5        2.96        51.9        3.0
      93.2        3.20        52.6        2.9
      93.6        2.90        51.7        5.5
      93.3        3.06        67.4        4.4
      94.7        3.36        70.0        4.1
      98.0        3.89        67.8        4.3
     100.7        3.79        60.9        6.8
     101.5        4.38        75.3        5.5
     103.1        4.41        74.8        5.5
     104.2        4.35        71.7        6.7
     105.4        4.33        83.0        5.5
     106.7        4.26        87.1        5.7
     108.1        4.40        94.0        5.2
     109.9        4.49       107.4        4.5
     113.1        5.13       118.0        3.8
/*
//GO.FT05F002 DD *
    2    4 4 1 2 3
/*
//
```

267

SAMPLE OUTPUT

```
TESTING REGAN1

ECHO CHECK
 NC.CF VARIABLES =          4
 NO.OF OBSERVATIONS =      21
 MS1                        1
 MS2                        2
 ER1                       0.0
 ER2                       0.0
 (4F10.0)

    68.000      2.530     30.600      3.800

ACTUAL NOB =   21

   VAR    MEAN              STD. DEV.          POP. S.D.
    1     0.954713E 02     0.115109E 02     0.112334E 02
    2     0.357190E 01     0.812613E 00     0.793029E 00
    3     0.663475E 02     0.229587E 02     0.224054E 02
    4     0.472857E 01     0.113981E 01     0.111234E 01


CORR. MATRIX   R(I, 1)        R(I, 2)        R(I, 3)        R(I, 4)        R(I, 5)        R(I, 6)        R(I, 7)



R( 2,J)         0.919635

R( 3,J)         0.915466        0.901840

R( 4,J)         0.342833        0.355816        0.164767

DEPENDENT VARIABLE IS NUMBER  4,   INDEPENDENT VARIABLES ARE  1, 2, 3,

    MLT.R.               STD.ERROR             F                DF
   0.56936973E 00     0.10163383E 01     0.27182302E 01     0.17000000E 02

      B                 BETA                 F
   0.75735092E-01     0.76484263E 00     0.16857405E 01
   0.10165939E 01     0.72476649E 00     0.17452154E 01
  -0.59031434E-01    -0.11890440E 01     0.49302368E 01

INTERCEPT       -0.22165442E 01


  TABLE OF RESIDUALS

     Y              Y           RESIDUAL      RESIDUAL
  ESTIMATE        VALUE                         STE
   3.69906       3.80000       0.10094       0.09931          1
   4.32189       3.80000      -0.52189      -0.51350          2
   4.28140       3.80000      -0.48140      -0.47367          3
   4.66619       5.90000       1.23381       1.21398          4
   3.59993       5.30000       1.70007       1.67274          5
   4.04438       3.30000      -0.74438      -0.73241          6
   4.73434       3.00000      -1.73434      -1.70646          7
   4.99001       2.90000      -2.09001      -2.05641          8
   4.76846       5.50000       0.73154       0.71978          9
   3.98160       4.40000       0.41840       0.41168         10
   4.23912       4.10000      -0.13912      -0.13689         11
   5.15771       4.30000      -0.85771      -0.84393         12
   5.66786       6.80000       1.13214       1.11394         13
   5.47818       5.50000       0.02182       0.02147         14
   5.65937       5.50000      -0.15937      -0.15681         15
   5.86468       6.70000       0.83532       0.82189         16
   5.26818       5.50000       0.23182       0.22810         17
   5.05344       5.70000       0.64656       0.63616         18
   4.89448       5.20000       0.30552       0.30061         19
   4.33127       4.50000       0.16873       0.16602         20
   4.59851       3.80000      -0.79851      -0.78567         21

0.17558823E 02
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $0.36.

Charge to user = computer costs + postage + network overhead
= $0.36 + postage + network overhead

CONTENTS—REGAN1

pages	
1	Identification & Abstract
3– 5	User Instructions
7– 8	I/O
9	Cost—Contents

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $0.50.

Charge to user = computer costs + postage + network overhead
= $0.50 + postage + network overhead

CONTENTS—REGAN1

pages	
1– 9	Identification & Abstract
11–13	User Instructions
15–16	I/O
17	Cost—Contents

DESCRIPTIVE TITLE	Analysis of Variance Using Between-Subject Designs
CALLING NAME	SABCA
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Dr. Charles Spiker Institute of Child Behavior The University of Iowa Converted for 360/65 by Ted Sjoerdsma The University of Iowa Computer Center
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

This analysis-of-variance routine uses between-subject designs for one, two, or three factors as described by Lindquist[1]. It includes the simple randomized, the two-dimensional factorial (AxB), and the three-dimensional factorial (AxBxC). (*Note*: It includes the 7044 programs SIMRAN, ABFT, and ABCFT.) Proportionality must exist between the corresponding cells (treatment sub-groups) from row to row or from column to column of the table for the AxB; for AxBxC there must be proportionality between the cells of at least two oblongs in the same layer and between corresponding oblongs for all layers. The program can handle up to 830 scores per treatment group and a maximum of 25 treatment groups in each of the A, B, and C dimensions. Input data may be on cards, tape or disk. Output includes a summary table as well as sums, means, variances, and standard deviations for each group. For the AxB and the AxBxC designs, the relevant marginal sums and means are also printed.

SABCA is a between-subject design for one (A), two (A and B), or three (A, B, and C) dimensions; that is, a completely randomized

continued

design concerning inter-subject relations using different subjects for each treatment group. The AxB design can be considered as b repetitions of the simple randomized design where b is the number of treatments in the B dimension. Likewise, AxBxC can be regarded as c repetitions of AxB. Thus AxB is treated as AxBxC with one C treatment, and the simple randomized is treated as AxBxC with one B and one C treatment. The program uses exactly the method described by Lindquist[1] (pp. 49-56, 121-125, Chs. 9 and 10).

Restrictions

Proportionality as described above is required. Otherwise, negative sums of squares may result.

The user should be sure that the sum of squared scores, as well as the other sums, is less than 10^{75}. Otherwise, overflow will result causing an error and incorrect output.

The program handles up to 830 scores per treatment group and a maximum of 25 treatments per dimension.

Machine Requirements

The program runs under 360 operating system and can handle card, disk, or tape input. It requires 19K bytes of core storage.

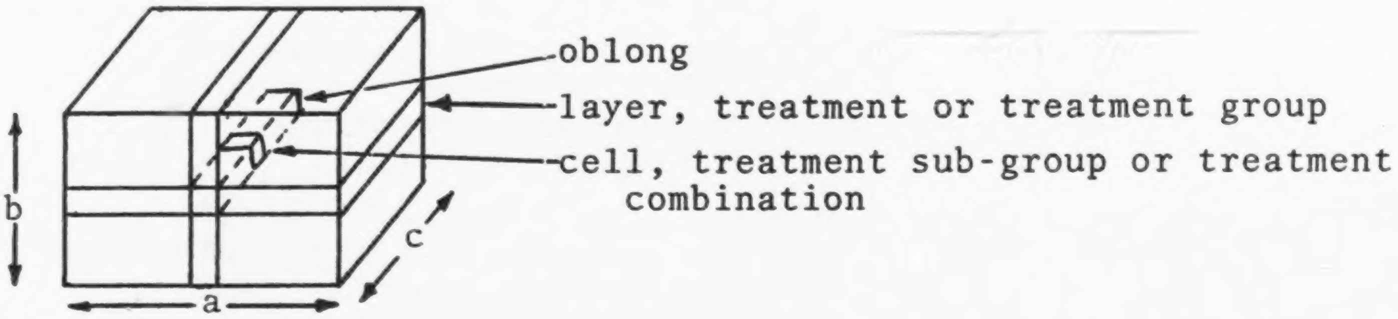
REFERENCES

1. Lindquist, E.F., *Design and Analysis of Experiments in Psychology and Education*, (Houghton Mifflin Company, Cambridge, Mass., 1953).

USER INSTRUCTIONS

Input

The input can be visualized as a three dimensional table AxBxC where a is the number of columns, b is the number of rows, and c is the number of slices.


oblong
layer, treatment or treatment group
cell, treatment sub-group or treatment combination
b
c
a

Thus if there are a columns, b rows, and c slices, there will be a + b + c treatments (or treatment groups) and abc treatment combinations (or treatment sub-groups). The data are read in row-by-row, slice-by-slice. That is, all groups in the first row of the first slice are read in, followed by all groups in the second row, and so on until all groups of the first slice have been read in. Then the groups of the first row of the second slice are read, followed by the second row of the second slice, etc. The procedure is continued until the last row of the last slice has been read. For the AxB this procedure simplifies to reading the data row-by-row, and for the simple randomized design the a treatment groups are read successively.

Data Control Cards

Data Control Cards are required in the following order.

Card	*Columns*	*Contents*
1	1–80	An E or F FORTRAN format enclosed in parentheses for the input score data.
2	1–80	The output sums format (E or F). One sum will be printed for each A treatment.
3	1–80	The output means format (E or F). One mean will be printed for each A treatment.
4	1–80	Title (alphameric characters which will be printed on the first page of the output).

continued

Card	*Columns*	*Contents*	
5	9–10	a: the number of A treatments	right-justified; a,b,c $\leq$ 25
	19–20	b: the number of B treatments (use a blank or a one for the simple randomized design)	
	29–30	c: the number of C treatments (use a blank or a one for the AxB and the simple randomized designs)	
	71	blank: card input of scores 1: tape or disk input of scores If tape or disk input is used, a DD card must be included for data set reference number 2.	

Subject and Data Cards

Subject Cards for one row of the slice should be as follows.

Columns	*Contents*	
1–10	Number of subjects for the first A treatment (or group) in the row and slice.	right justified
11–20	Number of subjects for second A treatment	
⋮	⋮	
61–70		

There can be 7 numbers per card; if there are more than 7 A-treatments, the number should continue in like manner on additional cards. Leading zeroes need not be punched. The maximum number of scores per treatment group is 830.

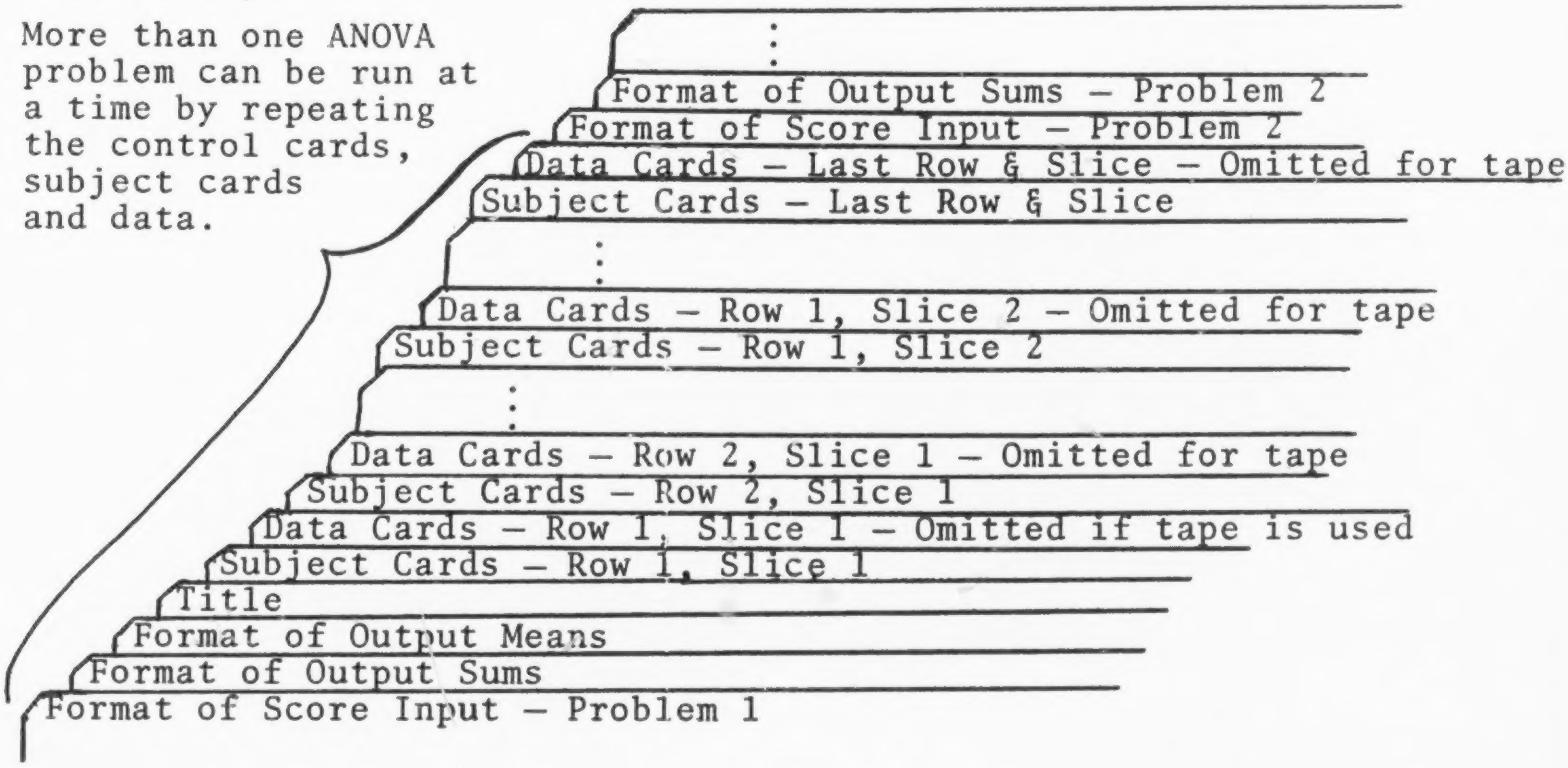
The Data Cards for the same row and slice follow immediately after the Subject Cards unless tape input is used. The scores are read according to the format specified on Data Control Card Number 1. All scores are read for the first group, then the second group, etc., as described above. Each card may contain only subject scores for one treatment (or treatment group).

The Subject and Data Card sets must be punched row by row for all rows.

continued

Deck Set-Up

More than one ANOVA problem can be run at a time by repeating the control cards, subject cards and data.


Format of Output Sums – Problem 2
Format of Score Input – Problem 2
Data Cards – Last Row & Slice – Omitted for tape
Subject Cards – Last Row & Slice
Data Cards – Row 1, Slice 2 – Omitted for tape
Subject Cards – Row 1, Slice 2
Data Cards – Row 2, Slice 1 – Omitted for tape
Subject Cards – Row 2, Slice 1
Data Cards – Row 1, Slice 1 – Omitted if tape is used
Subject Cards – Row 1, Slice 1
Title
Format of Output Means
Format of Output Sums
Format of Score Input – Problem 1

Output

The output includes a summary table containing degrees of freedom, sums of squares, mean squares, and F-ratios; sum of scores, sums of squared scores, means, sums of squared deviations, variances, and standard deviations for each group. For the AxB and the AxBxC designs, the relevant marginal sums and means are also printed.

Misuse and error messages are printed where appropriate.

SAMPLE INPUT

```
(7F10.0)
(7F10.0)
(7F10.2)
SAMPLE PROBLEM AXB FACTORIAL ANOVA SPIKER
         2         2         1
         8         8
        30        29        25        26        25        32        43
        29
        29        28        25        37        32        27        27
        31
         8         8
        27        36        25        24        28        21        31
        30
        19        25        24        24        24        39        23
        23
(16F5.1)
(5F12.3)
(5F12.3)
SAMPLE PROBLEM FOR SIMPLE RANDOMIZED DESIGN--PROB. 1 CHAP. 3 LINDQUIST--
         5         1         1
        15        15        15        15        15
15.0 18.0  9.0 11.0 13.0 20.0  9.0 13.0  5.0 10.0 22.0 18.0 17.0 10.0 12.0
16.0  8.0 12.0  5.0  9.0 10.0 12.0  8.0 11.0 18.0 12.0  8.0 11.0 10.0 12.0
 5.0 10.0 10.0  7.0 17.0 17.0 11.0  7.0  7.0  6.0  6.0  5.0  7.0  4.0  9.0
 9.0  8.0  8.0 14.0  5.0 11.0  9.0 16.0  7.0  4.0  9.0  8.0  8.0  8.0 10.0
 7.0  7.0  0.0  8.0 11.0  7.0  9.0  9.0  2.0  6.0  8.0  3.0  1.0  0.0  0.0
(5F10.0)
(5F10.0)
(5F10.2)
SAMPLE PROBLEM FROM LINDQUIST, PP. 116-117
         4         3         1
         5         5         5         5
        34        19        24        36        25
        16        35        18        16        12
        30        23        39        29        24
        38        32        21        36        15
         5         5         5         5
         7        24        12        43        41
        41        30        17        31        24
        28        34        40        27        42
        48        25        29        22        28
         5         5         5         5
        39        25        40        57        39
        19        49        30        46        28
        20        36        42        12        53
        30        23        24        38        37
(8F10.0)
(8F10.0)
(8F10.2)
```

continued

```
SAMPLE AXBXC FACTORIAL ANOVA PROBLEM
         3         4         2
         4         4         4
         9         7         7        14
        15        10        13        16
        10        10        15        15
         6         6         6
        14         7         5        15        12        11
         8         6        10        13        13        14
         9        13         7        13        12         8
         8         8         8
        18        16        12        16        10         9        11        12
        19        13         9        13         7        13         7         9
        20        17        13        16         9        14        14        12
         5         5         5
        16         9        17         9        22
        17        16        19        10        21
        17        11        15        12        14
         4         4         4
        13        12        18        18
        16        10        18        20
        16         9        14        19
         6         6         6
        17        22        14        17        10        12
         7        16         7        13        16         9
        15        23        16        15        11        13
         8         8         8
        29        25        18        25        15        21        18        30
        21        18        19        21        15        15        12        18
        19        24        15        22        17        17        11        24
         5         5         5
        30        21         9        13        18
        21        15         6        16        18
        15        16        17        23        17
/*
//
```

SAMPLE OUTPUT

```
SAMPLE PROBLEM AXB FACTORIAL ANOVA SPIKER                                  0359
         2   A-TREATMENTS,          2   B-TREATMENTS,          1   C-TREATMENTS.

NUMBER OF C-TREATMENTS IS ONE.  A BY B FACTORIAL IS ASSUMED .
SLICE  1      ROW  1
SUMS FOR CELLS
     239.      236.
SUMS OF SQUARED SCORES
    7381.     7062.
CELL MEANS
    29.88     29.50
CELL SUMS OF SQUARED DEVIATIONS
     241.      100.
CELL VARIANCES
    30.11     12.50
CELL STANDARD DEVIATIONS
     5.49      3.54
SLICE  1      ROW  2
SUMS FOR CELLS
     222.      201.
SUMS OF SQUARED SCORES
    6312.     5293.
CELL MEANS
    27.75     25.13
CELL SUMS OF SQUARED DEVIATIONS
     151.      243.
CELL VARIANCES
    18.94     30.36
CELL STANDARD DEVIATIONS
     4.35      5.51
MARGINAL STATISTICS, SLICE  1
COLUMN SLICE SUMS
     461.      437.
COLUMN SLICE MEANS
    28.81     27.31
ROW SLICE SUMS
     475.      423.
ROW SLICE MEANS
    29.69     26.44

SUMMARY TABLE

           DF           SUMS OF SQUARES       MEAN SQUARES    F-RATIOS

COLUMNS     1.           18.00000000         18.00000000      0.6855

ROWS        1.           84.50000000         84.50000000      3.2180

CXR         1.           10.12500000         10.12500000      0.3856

WITHIN     28.          735.25000000         26.25892857

TOTAL      31.          847.87500000         27.35080645

END OF ANALYSIS   1
```

continued

```
 SAMPLE PROBLEM FOR SIMPLE RANDOMIZED DESIGN--PROB. 1 CHAP. 3 LINDQUIST--
         5   A-TREATMENTS,          1   B-TREATMENTS,          1   C-TREATMENTS.

NUMBER OF B-TREATMENTS IS ONE. SIMPLE RANDOMIZED IS ASSUMED.
SLICE  1       ROW  1
SUMS FOR CELLS
   202.000     162.000     128.000     134.000      78.000
SUMS OF SQUARED SCORES
  3036.000    1896.000    1314.000    1326.000     608.000
CELL MEANS
    13.467      10.800       8.533       8.933       5.200
CELL SUMS OF SQUARED DEVIATIONS
   315.733     146.400     221.733     128.933     202.400
CELL VARIANCES
    21.049       9.760      14.782       8.596      13.493
CELL STANDARD DEVIATIONS
     4.588       3.124       3.845       2.932       3.673

SUMMARY TABLE

           DF          SUMS OF SQUARES       MEAN SQUARES   F-RATIOS

COLUMNS     4.         556.58666667         139.14666667     9.5944

WITHIN     70.        1015.20000000          14.50285714

TOTAL      74.        1571.78666667          21.24036036

END OF ANALYSIS   2
```

```
 SAMPLE PROBLEM FROM LINDQUIST, PP. 116-117
         4   A-TREATMENTS,          3   B-TREATMENTS,          1   C-TREATMENTS.

NUMBER OF C-TREATMENTS IS ONE.  A BY B FACTORIAL IS ASSUMED .
SLICE  1       ROW  1
SUMS FOR CELLS
    138.        97.       145.       142.
SUMS OF SQUARED SCORES
   4314.      2205.      4367.      4430.
CELL MEANS
   27.60      19.40      29.00      28.40
CELL SUMS OF SQUARED DEVIATIONS
    205.       323.       162.       397.
CELL VARIANCES
   41.04      64.64      32.40      79.44
CELL STANDARD DEVIATIONS
    6.41       8.04       5.69       8.91
SLICE  1       ROW  2
SUMS FOR CELLS
    127.       143.       171.       152.
SUMS OF SQUARED SCORES
   4299.      4407.      6033.      5038.
CELL MEANS
   25.40      28.60      34.20      30.40
CELL SUMS OF SQUARED DEVIATIONS
   1073.       317.       185.       417.
CELL VARIANCES
  214.64      63.44      36.96      83.44
CELL STANDARD DEVIATIONS
   14.65       7.96       6.08       9.13
SLICE  1       ROW  3
SUMS FOR CELLS
    200.       172.       163.       152.
SUMS OF SQUARED SCORES
   8516.      6562.      6413.      4818.
CELL MEANS
   40.00      34.40      32.60      30.40
CELL SUMS OF SQUARED DEVIATIONS
    516.       645.      1099.       197.
CELL VARIANCES
  103.20     129.04     219.84      39.44
CELL STANDARD DEVIATIONS
   10.16      11.36      14.83       6.28
```

continued

```
MARGINAL STATISTICS, SLICE  1
CCLUMN SLICE SUMS
    465.       412.       479.       446.
COLUMN SLICE MEANS
   31.00      27.47      31.93      29.73
ROW SLICE SUMS
    522.       593.       687.
ROW SLICE MEANS
   26.10      29.65      34.35
```

```
SUMMARY TABLE

            DF            SUMS OF SQUARES       MEAN SQUARES    F-RATIOS

COLUMNS      3.            168.33333333          56.11111111     0.4864

ROWS         2.            685.03333333         342.51666667     2.9689

CXR          6.            590.96666667          98.49444444     0.8538

WITHIN      48.           5537.60000000         115.36666667

TOTAL       59.           6981.93333333         118.33785311

ENC OF ANALYSIS   3
```

```
SAMPLE AXBXC FACTORIAL ANCVA PROBLEM
        3   A-TREATMENTS,          4   B-TREATMENTS,          2   C-TREATMENTS.
SLICE  1       ROW  1
SUMS FCR CELLS
     37.        54.        50.
SUMS OF SQUAREC SCORES
    375.       750.       650.
CELL MEANS
    9.25      13.50      12.50
CELL SUMS OF SQUARED DEVIATIONS
     33.        21.        25.
CELL VARIANCES
    8.19       5.25       6.25
CELL STANDARC CEVIATIONS
    2.86       2.29       2.50
SLICE  1       ROW  2
SUMS FCR CELLS
     64.        64.        62.
SUMS OF SQUARED SCORES
    760.       734.       676.
CELL MEANS
   10.67      10.67      10.33
CELL SUMS OF SQUARED DEVIATIONS
     77.        51.        35.
CELL VARIANCES
   12.89       8.56       5.89
CELL STANDARC DEVIATIONS
    3.59       2.92       2.43
SLICE  1       ROW  3
SUMS FCR CELLS
    104.        90.       115.
SUMS OF SQUARED SCORES
   1426.      1128.      1731.
CELL MEANS
   13.00      11.25      14.37
CELL SUMS OF SQUARED DEVIATIONS
     74.       115.        78.
CELL VARIANCES
    9.25      14.44       9.73
CELL STANDARD DEVIATIONS
    3.04       3.80       3.12
```

continued

```
SLICE  1      ROW  4
SUMS FOR CELLS
      73.        83.       69.
SUMS OF SQUARED SCORES
    1191.     1447.       975.
CELL MEANS
    14.60     16.60     13.80
CELL SUMS OF SQUARED DEVIATIONS
     125.        69.       23.
CELL VARIANCES
    25.04     13.84      4.56
CELL STANDARD DEVIATIONS
     5.00      3.72      2.14
MARGINAL STATISTICS, SLICE  1
COLUMN SLICE SUMS
     278.       291.      296.
COLUMN SLICE MEANS
    12.09     12.65     12.87
ROW SLICE SUMS
     141.       190.      309.       225.

ROW SLICE MEANS
    11.75     16.56     12.87     15.00
SLICE  2      ROW  1
SUMS FOR CELLS
      61.        64.       58.
SUMS OF SQUARED SCORES
     961.      1080.      894.
CELL MEANS
    15.25     16.00     14.50
CELL SUMS OF SQUARED DEVIATIONS
      31.        56.       53.
CELL VARIANCES
     7.69     14.00     13.25
CELL STANDARD DEVIATIONS
     2.77      3.74      3.64
SLICE  2      ROW  2
SUMS FOR CELLS
      92.        68.       93.
SUMS OF SQUARED SCORES
    1502.       860.     1525.
CELL MEANS
    15.33     11.33     15.50
CELL SUMS OF SQUARED DEVIATIONS
      91.        89.       84.
CELL VARIANCES
    15.22     14.89     13.92
CELL STANDARD DEVIATIONS
     3.90      3.86      3.73
SLICE  2      ROW  3
SUMS FOR CELLS
     181.       139.      149.
SUMS OF SQUARED SCORES
    4305.      2485.     2921.
CELL MEANS
    22.62     17.38     18.62
CELL SUMS OF SQUARED DEVIATIONS
     210.        70.      146.
CELL VARIANCES
    26.23      8.73     18.23
CELL STANDARD DEVIATIONS
     5.12      2.96      4.27
SLICE  2      ROW  4
SUMS FOR CELLS
      91.        76.       88.
SUMS OF SQUARED SCORES
    1915.      1282.     1588.
CELL MEANS
    18.20     15.20     17.60
CELL SUMS OF SQUARED DEVIATIONS
     259.       127.       39.
CELL VARIANCES
    51.76     25.36      7.84
CELL STANDARD DEVIATIONS
     7.19      5.04      2.80
```

continued

```
MARGINAL STATISTICS, SLICE  2
COLUMN SLICE SUMS
    425.      347.       388.
COLUMN SLICE MEANS
   18.48     15.09      16.87
ROW SLICE SUMS
    183.      253.       469.       255.

ROW SLICE MEANS
    15.25     14.06      19.54      17.00
MARGINAL STATISTICS  R1-C1, R1-C2 TO R 4-C 3
ROW COLUMN SUMS
      98.      118.       108.       156.       132.       155.       285.       229.
     264.      164.       159.       157.
ROW COLUMN MEANS
    12.25     14.75      13.50      13.00      11.00      12.92      17.81      14.31
    16.50     16.40      15.90      15.70
MARGINAL STATISTICS, CUBE
COLUMN SUMS
     703.      638.       684.
COLUMN MEANS
    15.28     13.87      14.87
ROW SUMS
     324.      443.       778.       480.
ROW MEANS
    13.50     12.31      16.21      16.00
SLICE SUMS
     865.     1160.
SLICE MEANS
    12.54     16.81

SUMMARY TABLE

               DF          SUMS OF SQUARES        MEAN SQUARES   F-RATIOS

COLUMNS      2.           48.56521739          24.28260870      1.3965

ROWS         3.          400.77053140         133.59017713      7.6851

SLICES       1.          630.61594203         630.61594203     36.2777

CXR          6.          109.79867150          18.29977858      1.0527

CXS          2.           91.31884058          45.65942029      2.6267

RXS          3.          116.46739130          38.82246377      2.2334

CXRXS        6.           67.12282609          11.18713768      0.6436

WITHIN     114.         1981.66666667          17.38304094      0.0

TOTAL      137.         3446.32608696          25.15566487      0.0

END OF ANALYSIS   4
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $0.56.

Charge to user = computer costs + postage + network overhead
= $0.56 + postage + network overhead

CONTENTS—SABCA

pages
1– 2 Identification & Abstract
3– 5 User Instructions
7–13 I/O
15 Cost—Contents

DESCRIPTIVE TITLE	Numerical Frequency Analysis
CALLING NAME	NUMFREQ
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine The University of Iowa Computer Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

NUMFREQ gives a frequency distribution table similar to those found in most statistical books. It tells how many times a value occurred in the variable and gives the percentage of the total, cumulative percentage, and cumulative frequency for each value. In addition, the mean, standard deviation and standard error of the mean are calculated for each variable. Missing data are represented as a blank or minus zero, or by using missing data codes.

USER INSTRUCTIONS

Input Deck

The input deck consists of the following cards in the order presented.

Title Card

Columns	*Contents*
1–72	Any alphanumeric information desired to identify the data and output.

Problem Card

Columns	*Contents*
1– 6	PROBLM
7–10	Number of variables (≤ 500).
11–14	Number of cases (≤ 2000).
15	The number of decimal places to be printed in the table. No more than 6 places are allowed to the right of the decimal point. Zero must be punched if no decimal places are desired.
16	Number of format cards (≤ 6).
17	1: Variable names are included. 0: Variable names are not included.
18	Number of missing data codes (≤ 3).
19	blank: card input 8 or 9: disk or tape input on unit 8 or 9. *NOTE:* DD Cards must be used to define this file. They must be placed before the //GO.SYSIN DD * card.

Variable Format Cards

Columns	*Contents*
1–72	The format statement may be continued for a maximum of 6 cards. The number of actual cards should be put in Col. 16 on the Problem Card. The standard E or F type FORTRAN statements are used. The format should be punched with no blanks between the beginning and ending parentheses.

continued

Missing Data Code Card (Included only if Col. 18 of the Problem Card is not 0)

Up to three codes are allowed.

Columns	*Contents*
1– 5	Fill in missing data codes in as many
6–10	of these spaces as indicated in Col. 18
11–15	of the Problem Card.

Variable Name Cards (Optional)

Variable Name Cards can be indicated in Col. 17 of the Problem Card. Twelve names are allowed per card using six columns each. If names are included, use enough cards to name all variables.

Data Cards

The data should be punched variable by case according to the format statement. The data for each case must start on a new card and may continue for as many cards as necessary. All cases must have an equal number of cards.

If more problems are to follow, they should be placed *before* the STOP Card. The Title Card through the Data Cards must be repeated. The STOP card is used only at the end of the complete run.

Stop Card

Columns	*Contents*
1– 4	STOP

Job Control

```
//name JOB (project no., etc.)
//step EXEC STAT,PARM.GO='NUMFREQ',REGION.GO=130K
//GO.SYSIN DD *
   input deck
/*
```

Methods

After the data are read into the machine one case at a time by variables, the matrix is transposed. Each variable is then scanned and the values are counted. Calculations of the mean, standard deviation and standard error are made.

000 0109

SAMPLE INPUT

```
LARGE TEST WITH MISSING DATA--NUMFREQ
PROBLM  31 1001110
(F3.0,1X,2F2.0,F3.0,2F2.0,F3.0,8F1.0,4F2.0,12F1.0)
CASENO PATNO  HOSP   SEX  RACE  DATE  YEARNOACMT  DIAGHEIGHTWEIGHTHGT/WT
FEVER1FEVER2FEVER3BLDPR1BLDPR2BLDPR3  URN1  URN2  URN3  EKG1  EKG2  EKG3
RESLT1RESLT2RESLT3 DRUG1 DRUG2 DRUG3  POST
529 12 6 54 4 4 3336253231 1144312433223101243121516
530 12 6 24 5 6 63212233233 155212253323 91353132612
533 13 6 14 610 532            43 12454422     631216
534X14 5 74 5 4 632683121452267412243524121843222623
536 14 6 24 611 73398311314226324  3531       1 121536
537 14 4 62 8    91215  231 2 2111  52 2320206 21123
538X12 6 54 1 7 6333222  1 2 53 1237322220206 222324
539 14 6 34 913 633584112332243312343424 91444122444
540 14 6 33 3 4 232453111131 2313  71 1      7 342625
542 14 6 13 1 1 134323111131231311 5352220  64132624
501 14 6 741017 523423122122124211 6342218  63132525
502 14 6 34 710 8336231111  142111 53422 6  53121623
503 14 6 34 1    144525321422164111 42425 72044122612
507 14 6 34 9 7 9244853  1 2162111 43523 1 534321545
508 13 6 74 4 2 2248853     1164 11 33222 3  3312214
509 11 6 14 4 4 423323213132227111 43224 21733121642
510X14 6 34 4 5 5223241323 1172132333325 41233241232
511 11 6 14 1     121           11111 73121     643117
513 13 6 24 6 9 423523112252115111 45528 31834121131
514 14 6 44 5 6 423825333132162111 26322 2  33122432
471 14 5 731118 823683312352162 12333424 51034132621
472 13 5 44 912 42232           62311 42325102044122636
473 14 4 63 6 8 61288533313117114  1321       1 121642
478X11 6 54 1 1 22352423 321143111 7112520206 222623
479 13 6 34 6 7 533323111132253511 53523141853322623
480 13 6 54 5 5 63332 3  1  131111 73 2220  6312263
481 13 6 14 1     334231111 1125111 63523202064122633
482 14 6 34 7 8 5336231311322622122243224 01833122533
486 13 3 51 5 5 523725311 3214513  42223192043122632
487 11 6 14 1     1452532212211113  56525202064132611
417 14 6 54 916 623323332152235211 36424 6 944221646
418 14 6 54 8 4 7236231  1 2172111 71128    6 321141
419 14 6 64 6 9 523623332132172112236523 7 944122634
420 13 6162 7 8 8421553     1 15112351 2317205332161
421 12 4 72 6 6 53332431323213313  2431       1 322533
422X14 6 14 1 1 213123233132232111 7211       1 231111
423 13 6 14 4 5  3312           1313  62323     63121612
427 14 5 23 8 9 724581          143111 44423101444122531
429 11 6 44 5 8 5235231121 2241111 43329 32033122611
430 12 6 34 5 5 4331231112 2 41212353323152053  2623
431X12 2 24 612 4231233221 1123112 32324 4 733221411
432 14 6 44 8 9 53482 2133 2 41512245423 81344122646
433 14 6 34 7 6 933225322121275111 36222 3  34122626
434 13 6 34 910 62212          231111 35527 41434212623
434X13 6 74 910 72212          231111 35527 41434222624
437 14 6 6410 2 83378411133214213  54424192064212433
438 14 6 34 911 633925312132166111 23423 4 533122612
439 14 4 42 911 8332241114 1135212435324 4 734141114
440 14 4 11 3 2 324223111152251 3  34323 1 53 131221
441 11 6 54       421           2111224432315194 122643
442 14 6 14 5 8 6225211  3   54212353221     63122621
443 13 5 11 1     3212311115114613  34221      2 122521
447 13 6 14 4 4 4331253331 2146112313223 0 133121321
450 14 6 54 5 7 3237232111 2156111 43325 92043122623
451 13 6 64 5 7 42375           154111 42321      23312621
452 14 6 94 4 4 533285332451237111 13222 1  33122611
453 14 5 63 7 5 623753111452163132353223202063121525
454 13 6 54 2   3222232   11 311   7111      7 122641
454X12 6 24 5   4226232  1  27523  6432320206 222642
457 14 6 641313 72358321 22216311236122      5 121224
458 14 6 24 610 523583311132161612334324 5 933121625
459 14 6 24 912 63338123 2 113611235442211  53132513
460 14 6 52 2    223 231     2 23211 7612220  631222 4
461 14 6 84       332           31111 23323 3 533122523
462X14 6 34 6 9 6336 53 23 2154211 33325 2 734222722
```

continued

000 0109

```
463 14 6 24 510 533623111152175111  33323 2 933121631
464 13 6 71 1     222253211321233                 122636
466 14 6 24 7 7 4227532222332273312342527 41534132643
467 14 5 54 8 8 734825313332163112333324 3 734121421
469 13 6 24 2 2 234525312122154112253423202064322621
488 12 6 53 7 9 333324111232122212244324151843121612
489 14 6 44 4 6 4236231131 2164111 5322220  63122646
491 13 6 74 8 7 633654312332256112353324132053122622
492 12 6 44 2 1 333524231332223112353325 92053122611
493 13 6 14 5 2 42382411331216711235242220  64122623
496 13 6 84 5 6 322325312122154112322325 3 633131521
497 14 6 24 5 8 523622221232276512356424122054132427
498 14 6 24 1 2 223123332132145 11 33223 3 433132631
499 13 6 24 5 7 5328253111321433122345 21     24122625
500 12 6 54 3 2 5337231113 215613  51 25202063241211
516X13 6 44 8 5 633285331132176112214221      2 231334
517 13 6 24 711 73332433345215324  45421      23121625
518 13 6 34 916 922223223152127111 43222 9  44122633
520 14 6 33 913 7345831221 2221411 43525 81644122624
521 13 6 24 910 632323112132147212343529 72044122623
522 13 6 14 5 6 624653232132167 11 36324 21133122621
524 12 5 31 2 1 2233212  15222344  44424 7124 122524
526 12 6 51 3 5 323223311121243 2  5442320206 122611
527 13 6 34 2 3 123645311132161111 33422 3  34122625
528 14 4 12 4 5 52342533313114712  2321      1 122531
543 14 6 44 2 3 33272411332216413  6311520206 132611
544X14 6 74 814 933984131 2217623  6122420206 221245
545 13 6 13 810 624425312132143112233323 0 234122613
545A13 6 54 8 5 723584112 32256112243325 91643121222
546 12 5 14 3 4 23412412123213114  4541      1   2611
547 12 6 31 9 8 6331531111 1131311 7442220  6 122624
549 14 6 24 5 6 433685331122167111 33321      23  2624
550 14 6 34 5 9 632823111132166111 33423 1 634152624
551 14 6 54 914 9336833111 216211222 44 23 1 434132432
552 13 6 24 5 6 233425313122142 12233222 3  33131541
STOP
/*
//
```

SAMPLE OUTPUT

```
UNIVERSITY OF IOWA COMPUTER CENTER
 FREQUENCY ANALYSIS PROGRAM

LARGE TEST WITH MISSING DATA--NUMFREQ

NUMBER OF VARIABLES =   31       NUMBER OF CASES =  100




VARIABLE NUMBER     1     CASENO

OBSERVED                CUMULATIVE               CUMULATIVE
 VALUE      FREQUENCY   FREQUENCY      PERCENT     PERCENT

   417.0          1            1         1.00        1.00
   418.0          1            2         1.00        2.00
   419.0          1            3         1.00        3.00
   420.0          1            4         1.00        4.00
   421.0          1            5         1.00        5.00
   422.0          1            6         1.00        6.00
   423.0          1            7         1.00        7.00
   427.0          1            8         1.00        8.00
   429.0          1            9         1.00        9.00
   430.0          1           10         1.00       10.00
   431.0          1           11         1.00       11.00
   432.0          1           12         1.00       12.00
   433.0          1           13         1.00       13.00
   434.0          2           15         2.00       15.00
     .                         .                      .

     .                         .                      .

     .                         .                      .
   521.0          1           75         1.00       75.00
   522.0          1           76         1.00       76.00
   524.0          1           77         1.00       77.00
   526.0          1           78         1.00       78.00
   527.0          1           79         1.00       79.00
   528.0          1           80         1.00       80.00
   529.0          1           81         1.00       81.00
   530.0          1           82         1.00       82.00
   533.0          1           83         1.00       83.00
   534.0          1           84         1.00       84.00
   536.0          1           85         1.00       85.00
   537.0          1           86         1.00       86.00
   538.0          1           87         1.00       87.00
   539.0          1           88         1.00       88.00
   540.0          1           89         1.00       89.00
   542.0          1           90         1.00       90.00
   543.0          1           91         1.00       91.00
   544.0          1           92         1.00       92.00
   545.0          2           94         2.00       94.00
   546.0          1           95         1.00       95.00
   547.0          1           96         1.00       96.00
   549.0          1           97         1.00       97.00
   550.0          1           98         1.00       98.00
   551.0          1           99         1.00       99.00
   552.0          1          100         1.00      100.00

 MEAN =    0.485760000000D 03

 STD. DEV. =    0.409651539529D 02

 STD. ERROR =    0.409651539529D 01
```

continued

```
VARIABLE NUMBER     2        PATNO

OBSERVED                  CUMULATIVE                   CUMULATIVE
 VALUE       FREQUENCY    FREQUENCY       PERCENT        PERCENT

   11.0            6              6          6.00          6.00
   12.0           14             20         14.00         20.00
   13.0           31             51         31.00         51.00
   14.0           49            100         49.00        100.00

MEAN =    0.1323000000000D 02

STD. DEV. =    0.9084897001910 00

STD. ERROR =    0.9084897001910-01

        .                                       .
        .                                       .
        .                                       .


VARIABLE NUMBER     5        RACE

OBSERVED                  CUMULATIVE                   CUMULATIVE
 VALUE       FREQUENCY    FREQUENCY       PERCENT        PERCENT

    1.0           11             11         11.22         11.22
    2.0            7             18          7.14         18.37
    3.0            5             23          5.10         23.47
    4.0            9             32          9.18         32.65
    5.0           20             52         20.41         53.06
    6.0           10             62         10.20         63.27
    7.0            8             70          8.16         71.43
    8.0           10             80         10.20         81.63
    9.0           14             94         14.29         95.92
   10.0            2             96          2.04         97.96
   11.0            1             97          1.02         98.98
   13.0            1             98          1.02        100.00

MEAN =    0.5530612244900 01

STD. DEV. =    0.2773973597240 01

STD. ERROR =    0.2802136704450 00

DATA ARE MISSING FOR     2 CASES.

        .                                       .
        .                                       .
        .                                       .


VARIABLE NUMBER    31        POST

OBSERVED                  CUMULATIVE                   CUMULATIVE
 VALUE       FREQUENCY    FREQUENCY       PERCENT        PERCENT

    3.0           43             43         56.58         56.58
    4.0           33             76         43.42        100.00

MEAN =    0.3434210526320 01

STD. DEV. =    0.4989462580490 00

STD. ERROR =    0.5723306496590-01

DATA ARE MISSING FOR    24 CASES.
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $1.56.

Charge to user = computer costs + postage + network overhead
= $1.56 + postage + network overhead

CONTENTS—NUMFREQ

pages
1 Identification & Abstract
3– 4 User Instructions
5– 8 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Simple Multiple Linear Regression
CALLING NAME	MISREGN
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine University Computer Center
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

MISREGN performs multiple linear regression, producing means, standard deviations and correlation coefficients. The multiple R, standard error of estimates, F ratio, degrees of freedom of regression, Beta weights, regression coefficients and y intercept are calculated from a chosen subset of the correlation matrix. Residuals may be computed. Correlations are calculated using a pairwise technique. Fifty variables are allowed. Missing data are allowed in the form of blanks or minus zero (-0).

Caution: The distortion due to the missing values may cause the results of multivariate analysis to be meaningless.

REFERENCES

Cooley, W.W., and Lohnes, P.R., *Multivariate Procedures for the Behavioral Sciences*, (John Wiley & Sons, Inc., N.Y., 1962), pp. 31-35.

USER INSTRUCTIONS

Input Deck

The input deck consists of the sets of cards in the order described.

Title Card

There is a single Title Card containing up to 80 columns of alphanumeric information to be printed as a heading for the output.

Data Control Card

Columns	*Contents*
1– 3	No. of variables ($\leq$ 50)
4	5: input from cards 2,3,4,8, or 9: unit number if data input from tape or disc.
5– 9	Number of observations ($\leq$ 99,999)
12	Number of Variable Format Cards to be read ($\leq$ 5)
14	0: Variable Name Cards are not read 1: Variable Name Cards are read

Variable Name Cards (optional; 12 names to a card)

Columns	*Contents*
1– 6	First variable name
7–12	Second variable name
⋮	⋮
67–72	And so on until *all* variables are named. Use as many cards as necessary.

Variable Format Card

Use Col. 1–72 only, starting with a left parenthesis and ending with a right parenthesis. An E or F format is allowed.

Data Cards

Regression Cards

One Regression Card must be present for each problem.

continued

000 0110

Regression Card (cont.)

Columns	*Contents*
5	1: residuals are not desired. 2: residual table is desired.
9–10	Number of variables to be used in regression (independent plus dependent)
11–12	Dependent variable number. (Variables are assigned numbers according to the order in which they are read; first variable read in is number one.
13–14	Independent variable numbers.
15–16	If there are more than 34 independent variables, continue on another card beginning with Col. 1–2.
⋮	⋮
79–80	

Machine Requirements

The MISREGN program uses a special OS Assembler Language program, BLNK.

SAMPLE INPUT

```
TESTING MISREGN ON MOONS DATA
  45   21  1
(4F10.0)
      68.0      2.53      30.6       3.8
      77.8      2.61      34.0       3.8
      83.8      2.82      46.0       3.8
      83.0      2.66      35.7       5.9
      83.8      2.62      54.1       5.3
      90.5      2.86      59.3       3.3
      92.5      2.96      51.9       3.0
      93.2      3.20      52.6       2.9
      93.6      2.90      51.7       5.5
      93.3      3.06      67.4       4.4
      94.7      3.36      70.0       4.1
      98.0      3.89      67.8       4.3
     100.7      3.79      60.9       6.8
     101.5      4.38      75.3       5.5
     103.1      4.41      74.8       5.5
     104.2      4.35      71.7       6.7
     105.4      4.33      83.0       5.5
     108.1      4.40      94.0       5.2
     106.7      4.26      87.1       5.7
     109.9      4.49     107.4       4.5
     113.1      5.13     118.0       3.8
    2    4 4 1 2 3
/*
//
```

SAMPLE OUTPUT

```
TESTING MISREGN ON MOONS DATA

          VARIABLE            NO. OF OBSERVATIONS            MEANS                   STANDARD DEVIATIONS


          1 - 1                        21                0.95471344E 02               0.11510864E 02
          2 - 2                        21                0.35718985E 01               0.81261265E 00
          3 - 3                        21                0.66347549E 02               0.22958649E 02
          4 - 4                        21                0.47285662E 01               0.11398106E 01



CORRELATIONS AND THE NUMBER OF OBSERVATIONS USED

              1       2       3       4

  1         1.0000
           (   21)

  2         0.9196  1.0000
           (   21) (   21)

  3         0.9155  0.9018  1.0000
           (   21) (   21) (   21)

  4         0.3428  0.3558  0.1648  1.0000
           (   21) (   21) (   21) (   21)



DEPENDENT VARIABLE IS NUMBER  4,   INDEPENDENT VARIABLES ARE  1, 2, 3,

    MLT.R.            STD.ERROR             F                DF
  C.56937599E 00    0.10163326E 01    0.27183180E 01    0.17000000E 02

      B               BETA               F
  0.75737655E-01    0.76486969E 00    0.20825405E 01
  0.10166178E 01    0.72478443E 00    0.21559839E 01
 -C.59033711E-01   -0.11890879E 01    0.60906029E 01

INTERCEPT      -0.22167225E 01



   TABLE OF RESIDUALS

       Y            Y         RESIDUAL     RESIDUAL
    ESTIMATE      VALUE                      STE
    3.69905      3.80000      0.10095      0.09933        1
    4.32185      3.80000     -0.52189     -0.51351        2
    4.28140      3.80000     -0.48140     -0.47367        3
    4.66620      5.90000      1.23380      1.21397        4
    3.59991      5.30000      1.70009      1.67277        5
    4.04436      3.30000     -0.74436     -0.73240        6
    4.73435      3.00000     -1.73435     -1.70648        7
    4.99003      2.90000     -2.09003     -2.05644        8
    4.76847      5.50000      0.73153      0.71977        9
    3.98158      4.40000      0.41842      0.41170       10
    4.23911      4.10000     -0.13911     -0.13687       11
    5.15773      4.30000     -0.85773     -0.84394       12
    5.66789      6.80000      1.13211      1.11392       13
    5.47820      5.50000      0.02180      0.02145       14
    5.65939      5.50000     -0.15939     -0.15683       15
    5.86471      6.70000      0.83529      0.82187       16
    5.26818      5.50000      0.23182      0.22809       17
    4.89447      5.20000      0.30553      0.30062       18
    5.05344      5.70000      0.64656      0.63617       19
    4.33124      4.50000      0.16876      0.16605       20
    4.59848      3.80000     -0.79848     -0.78564       21

0.17558792E 02
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $0.45.

Charge to user = computer costs + postage + network overhead
= $0.45 + postage + network overhead

CONTENTS—MISREGN

pages

1	Identification & Abstract
3– 4	User Instructions
5– 6	I/O
7	Cost—Contents

DESCRIPTIVE TITLE	Test-Scoring and Item-Analysis Package
CALLING NAME	(a) QUICKSCORE; (b) ITEMSTEP; (c) ITEMRS
INSTALLATION NAME	Wharton Computational Services Wharton School of Finance and Commerce University of Pennsylvania
AUTHOR(S) AND AFFILIATION(S)	Daniel Ashler Daniel Bricklin David Sheinson Wharton Computational Services
LANGUAGE	FORTRAN
COMPUTER	IBM 360/75
PROGRAM AVAILABILITY	Available to process data at Wharton Computational Services
CONTACT	Daniel Ashler, Wharton Computational Services, Wharton School of Finance and Commerce, Dietrich Hall, University of Pennsylvania, Philadelphic, Pa. 19104 Tel.: (215) 594-6422

FUNCTIONAL ABSTRACT

The Test-Scoring and Item-Analysis Package currently consists of three programs, QUICKSCORE, ITEMSTEP, and ITEMRS. Together, they provide a flexible facility for tasks that range from simple scoring of tests to the most sophisticated test analysis currently available. The best features of item-analysis programs in use on various campuses have been incorporated.

QUICKSCORE

...is the least expensive of the three programs to use. It scores tests and lists the examinees and their scores, first in alphabetical order and then in order of score. Up to 500 examinees can be scored at once. However, the input can be "batched;" that is, several groups of cards of up to 500 each can be scored in this manner, one after another. Beside each examinee's name is printed his Social Security number, his score (which has been corrected

continued

for chance success), the number of items that he got correct, that he got incorrect, that he omitted, and the number of items not reached. The corrected score gives +1 for each item answered correctly and $-1/(c-1)$ for each answered incorrectly, where c is the number of choices per item. Omitted items are scored 0.

A brief item analysis is then performed for each question. First, the estimated fraction of the examinees who knew the right answer is computed (difficulty of question). The examinees are then partitioned into an upper and a lower group on the basis of total test score, these groups being equal or nearly equal in size. A 2X2 table is constructed as follows.

	No. Rights	No. Wrongs
Upper half	A	B
Lower half	C	D

Two indices are computed on this table: Ø correlation coefficient (PHI) and a ratio index (RI). For a test with N choices per question and U examinees in the upper half and L examinees in the lower, the two indices are defined as follows.

$$\text{PHI} = \frac{AD - BC}{[(A+B)\ (A+C)\ (C+D)\ (B+D)]^{\frac{1}{2}}}$$

$$\text{RI} = \frac{2}{\pi} \arctan \log_2 (\text{UPSC}/\text{LOWSC}),$$

where

$$\text{UPSC} = \frac{A - B/(N-1)}{U} + \frac{1}{N-1}$$

and

$$\text{LOWSC} = \frac{C - D/(N-1)}{L} + \frac{1}{N-1}\ .$$

Both indices can have values in the range -1 to 1. The ratio index reflects the extent of discrimination between the upper and lower groups as well as the correlation of the score on the given question with total test score.

A vertical histogram of the score distribution follows, in which, however, minus scores, scores of zero, and scores of one are all counted in the score category of one. Test statistics follow, in which are given the number of questions in the test, the number of choices per item, the number of examinees, a copy of the scoring key, and the means, standard deviations, and variances of the scores.

continued

ITEMSTEP

...not only performs the same scoring as QUICKSCORE but also provides fuller item analysis and a stepwise-reduction feature. After scoring and analysis, one or more questions (items) are deleted on the basis of some criterion (supplied as a parameter), after which the shortened test is rescored and reanalyzed. The cycle of shortening and reprocessing is repeated as many times as specified. One of three different indices may be specified, on the basis of which the less desirable items shall be eliminated. The first is the W index, which is a weighted combination of the Davis[1] difficulty and discrimination indices. The second is the P-adjusted index, which is the proportion of correct responses, adjusted or corrected for chance success. The third is the contribution of the item to the total test variance. The user specifies either the index level required for retention of an item or that some fixed number of items with the worst index values be dropped at each cycle. A reliability coefficient is also computed. The stepwise-reduction process tends to produce successive score distributions that are more and more rectangular.

ITEMRS

...is the most advanced item-analysis program presently available. First, it provides a header page, in which the indices are explained. Then there are optional printouts of the main matrices produced during the run: namely, the sum-of-products matrix, variance–covariance matrix, and the product–moment correlation matrix. (It should be noted that the correlations are product–moment and not point biserial because the items are not scored dichotomously.) Next, an optional cluster analysis is performed on the correlations after they have been transformed to Fisher's Z's. (A cluster analysis is quicker and less expensive than factor analysis and for many purposes is just about as useful.) Each pass produces larger clusters and fewer of them.

Next, the examinee's scores are listed in a tabulation like that of QUICKSCORE, in order by name and then by score. Alongside the listing of the scores are 16 blank columns. If desired, data may be printed in these columns, such as scaled scores, grades, or comments. (For example, to print scaled scores, one provides a card for each possible score that can be obtained, specifying what shall be printed for that score.)

A large number of indices are computed for each item, including P (proportion of correct responses) and Q (proportion of incorrect responses), both unadjusted and adjusted for chance success; the variance contribution of the item; the product–moment correlation of the item with the total score; the "latent" correlation, which

continued

is the estimate of the correlation between the underlying ability to answer the item correctly and the total score; and the Fisher's Z transform of the product-moment correlation. Also, the effective number of choices (that is, the number of choices that actually attracted examinees) the estimated proportion who knew the answer (based on the effective number of choices), the Davis difficulty and discrimination indices, the W index, which is a weighted combination of the Davis indices, and the index DICAP, which measures the deviation of the responses to this item from the ideal pattern in which 50% of the examinees would choose the correct alternative and the remaining 50% would distribute equally over the remaining alternatives. Optionally, a tabulation of the number of responses to each alternative is given for each fifth of the examinees.

Following the item analysis, a histogram is printed, followed by test statistics, including the covariance reliability coefficient, the KR-20 reliability coefficient, the standard error of measurement, the variance error of measurement, and the test mean, standard deviation, and variance of total test scores.

REFERENCES

1. Davis, F.B., *Item Analysis Data* (Harvard University Press, Cambridge, Mass., 1949).

USER INSTRUCTIONS—QUICKSCORE

Description of Input

Title Card

Columns	*Contents*
2–80	Any information that the user wishes will be printed at the top of each page of output. Col. 1 should be blank.

Key Card

Columns	*Contents*
1– 3	KEY
4– n	Blanks (where n is at choice of user)
(n+1)	Sequence character, J on the first Key Card, K on the second, etc.
(n+2)–80	Response field. Each column corresponds to one test item, and is coded 0: 1st response 1: 2nd response : : 9: 10th response

The program will scan the response field and use the highest response code found to determine the number of choices per item. This number, in turn, determines the guessing correction used in scoring all the items. Therefore, it is important that, in a test with C choices per item, at least one item be keyed with the C^{th} choice. If any column is left blank, the corresponding test item will be omitted from the scoring and analysis.

Data Sheets or Cards

Data are generally submitted in the form of answer sheets. They may also be submitted in card form as follows, with one packet of cards per examinee, each packet consisting of as many cards as there are Key Cards, and organized similarly.

Columns	*Contents*
1– n	Identification field, identical on all cards of a packet. n corresponds to n on the Key Card
(n+1)	Sequence character

continued

Data Sheets or Cards *(continued)*

Columns	*Contents*
(n+2)–80	Examinee responses. Blank response columns will be considered as omitted items.

Description of Output

See Sample Output—QUICKSCORE, page 12.

USER INSTRUCTIONS—ITEMSTEP

Description of Input Data

Input data for ITEMSTEP are the same Title, Key, and Data Cards as for QUICKSCORE, followed by an End Card and some Criterion Cards.

End Card

Columns	*Contents*
1– 4	>END

Criterion Cards

There may be one or more cards in any of the following forms.

Fixed-Index Card

Columns	*Contents*
1–25	REMOVE ALL QUESTIONS WITH
27–29	aaa aaa is any legal I-, F-, or E-format number, such as 1.5.
31–80	VARIANCE. or WORTH. or P ADJ.

In the alternative phrases for Cols. 31–80, the final period is required and must not be separated from the word that it follows by any blanks.

Fixed-Questions Card

Columns	*Contents*
1–10	REMOVE THE
12–13	nn nn, any integer, is the number of questions to remove
15–80	QUESTIONS WITH THE LOWEST PADJ. or QUESTIONS WITH THE LOWEST WORTH. or QUESTIONS WITH THE LOWEST VAR.

continued

Passes Card(s)—First Card

Columns	*Contents*
1– 2	DO
4– 5	mm mm, any integer, is the number of passes to do
7–20	PASSES AND USE
22–26	P ADJ or WORTH or VAR
28–57	AS THE CRITERION FOR DROPPING.

Passes Card(s)—Second Card

Columns	*Contents*
1– 6	REMOVE
8– 9	qq qq, any integer, is the number of questions to remove
11–48	QUESTIONS IN EACH STEP WITH THE LOWEST
50 ff.	P ADJ. or WORTH. or VAR.

Description of Output

See Sample Output—ITEMSTEP, page 19.

USER INSTRUCTIONS—ITEMRS

Description of Input Data

Input data for ITEMRS are the same as for QUICKSCORE, with the exception that one to four Option Cards may be inserted after the Title Card and before the Key Cards. Each Option Card calls for a printout or punchout, which is otherwise suppressed.

Option Card(s)

Columns	*Contents*
1–80	MATRIX or CLUSTER or ITEM ANALYSIS or PUNCH

The MATRIX option calls for printout of the sum-of-products matrix, the variance—covariance matrix, and the product—moment correlation matrix. The CLUSTER option calls for a cluster analysis performed on the correlations after they have been transformed to Fisher's Z's. The ITEM ANALYSIS option calls for a large number of indices to be computed for each item. (See Sample Output.) The PUNCH option causes a card to be punched for each examinee, with his identification, score, number right, number wrong, number omitted, and number not reached.

Description of Output

See Sample Output—ITEMRS, page 29

SAMPLE INPUT

Since the input consists primarily of large numbers of Data Cards or answer sheets, it is not reproduced here.

SAMPLE OUTPUT—QUICKSCORE

```
+------------------------------------------------------------------------+
|                                                                        |
|                 T E S T   R U N   F O R   Q U I C K S C O R E          |
|                                                                        |
+------------------------------------------------------------------------+
```

EXPLANATION OF OUTPUT:

THE STUDENTS' NAMES AND SCORES ARE LISTED TWICE. THE FIRST LISTING IS IN ASCENDING ALPHABETICAL ORDER BY STUDENTS' NAMES. THE SECOND IS IN ASCENDING ORDER BY SCORE. BOTH OF THESE LISTINGS CONTAIN THE STUDENT'S NAME, HIS SOCIAL-SECURITY NUMBER, HIS TOTAL TEST SCORE (CORRECTED FOR GUESSING), THE NUMBER OF QUESTIONS TO WHICH HE RESPONDED CORRECTLY, THE NUMBER TO WHICH HE RESPONDED INCORRECTLY, THE NUMBER HE OMITTED, AND THE NUMBER HE DID NOT READ (E.G., IF HE OMITTED THE LAST FOUR QUESTIONS, THEY ARE COUNTED AS NON-READS.)

A BRIEF ITEM ANALYSIS IS PERFORMED ON EACH QUESTION. FIRST, THE FRACTION OF THE EXAMINEES WHO KNEW THE RIGHT ANSWER IS COMPUTED (DIFFICULTY). IN ADDITION, THE EXAMINEES ARE PARTITIONED INTO AN UPPER AND A LOWER GROUP ON THE BASIS OF TOTAL TEST SCORE. THESE GROUPS ARE EQUAL OR NEARLY EQUAL IN SIZE. FOR A GIVEN ITEM, THE APPROXIMATE PROPORTION OF THE UPPER GROUP WHO GOT IT RIGHT DIVIDED BY THE APPROXIMATE PROPORTION OF THE LOWER GROUP WHO GOT IT RIGHT YIELDS A RATIO, WHICH IS SUBJECTED TO A MONOTONIC TRANSFORMATION TO CONVERT IT INTO THE RATIO INDEX, WHICH CAN HAVE VALUES FROM -1 TO +1. THE RATIO INDEX REFLECTS THE CORRELATION OF ITEM SCORE WITH TOTAL TEST SCORE. THE PHI CORRELATION TEST IS ALSO PERFORMED.

AFTER THE ITEM ANALYSIS THERE IS A HISTOGRAM OF THE DISTRIBUTION OF SCORES. ALL SCORES ARE ROUNDED TO THE NEAREST INTEGER AND SCORES OF ZERO AND BELOW ARE TABULATED AS A SCORE OF ONE. EACH ASTERISK STANDS FOR ONE STUDENT. FINALLY, A SUMMARY OF THE MAJOR TEST STATISTICS IS PRINTED.

THIS OUTPUT WAS PRODUCED BY QUICKSCORE IV WRITTEN DURING JULY 1969 BY DANIEL BRICKLIN AND MODIFIED DURING NOVEMBER 1969 BY DAVID SHEINSON. IT IS THE PROPERTY OF WHARTON COMPUTATIONAL SERVICES OF THE UNIVERSITY OF PENNSYLVANIA.

continued

```
                                             T E S T   R U N   F O R    Q U I C K S C O R E
          NAME                                        SCORE            RIGHT     WRONG      OMIT  NON-READ

 LOMQUIST    ROBERTF      5- K-  J5                   50.75               56        21         3         0
ALTMAN       JEROLDD      5- K-  J5                   56.00               58         8         2        12
ASLANIAN     GREGORP      5- K-                       51.25               55        15         1         9
BAGLEY       BRUCE D      5- K-   5                   49.00               51         8         2        10
BALL OU      ROGER H      5-  -  D5                   52.50               55        10         2        13

BASS         ARTHUR       5-  -  E5                   38.50               44        22         9         5
BEAUCHAMP    GEOFFRL      5- K-   5                   51.00               55        16         0         9
BECHARA      DENNIS        - K-  L                    34.75               40        21         4        15
BEFFA        RICHA H      5- K-  I5                   27.00               34        28         1        17
BELGER       LARRY A      5- K-  G5                   41.50               46        18         1        15

BENKIN       RICHARL       - K-  J5                   47.75               53        21         1         5
BERGER       STEVENA      5- K-  B2                   27.50               32        18        15        15
BERSHAD      RUSSELB       - K-  I5                   37.75               43        21        11         5
BIALLAS      MIKE  D      5- K-  H5                   39.75               45        21         2        12
BLOOM        GARY         5- K-  G5                   47.25               49         7        19         5

BOONIN       DAVID M      5- K-  E5                   27.50               35        30         0        15
BRADEORD     CHARLEC      5- K-  H                    41.00               46        20        14         0
BRILL        JOEL  D       - C-  G5                   41.00               47        24         8         1
BRINN        STUART       5- K-  B5                   33.75               38        17         1        24
BUTLER       KENT  S      5- K-  A5                   38.25               45        27         0         8

   .                        .                                                      .
   .                        .                                                      .
   .                        .                                                      .

FULOP        RONALDG      5- K-  E5                   43.25               48        19        10         3
GARLAND      KIM   C      5- K-  A5                   44.50               47        10        23         0
GERSHON      STEVE A      5- K-  A5                   41.00               47        24         3         6
GIBBONS      JAMES T      5- K-  K5                   52.00               56        16         8         0
GLICKMAN     RONALDD      5- K-  G5                   33.25               38        19         1        22

GOODMAN      ROBERTA      5- K-  J5                   39.75               43        13         0        24
GREEN        JOHN  J      5- K-  I5                   52.50               57        18         2         3
GREENBAUM    KEITH R      5- K-  D5                   32.50               36        14         0        30
GREIF        STEVEND      5- K-  J2                   49.75               52         9         0        19
GRELLER      HOWARD       5- K-  B5                   49.00               53        16         0        11

HALL         KEITH R      5- K-  H5                   32.75               39        25         4        12
HALL         ROBERTL      5- K-  D5                   55.00               57         8         0        15
HALLAS       RICHARD      5- K-  B5                   41.00               45        16        18         1
JEZ          A BERTF      5- K-  E5                   27.50               35        30         9         6
LANGE        STEVE E      5-  -  D5                   57.50               59         6         0        15

LEGGETT      LAWRENE      5-  -  J5                   31.50               36        18         9        17
LEIGHTON     THOMASK      5-  -  B5                   24.75               31        25         0        24
LIBERI       KAREN A      5-  -  J5                   43.75               50        25         5         0
LIPMAN       DAVID J      5-  -  B5                   52.00               56        16         8         0
MARITZ       JAMES A      5-  -  B                    35.25               39        15         7        19

MARKS        MICHAEH      5- K-  K1                   71.25               73         7         0         0
MASSIE       DAVID W      5- K-  F5                   44.50               48        14        13         5
MAULE        JAMES E      5-  -  K5                   61.25               65        15         0         0
MCGINTY      STEPHEJ      5-  -  K5                   52.50               58        22         0         0
```

continued

308

```
                                         T E S T   R U N   F O R   Q U I C K S C O R E
        NAME                                         SCORE            RIGHT     WRONG      OMIT  NON-READ

LEIGHTON      THOMASK          5-  -  B5             24.75               31        25         0        24
BEFFA         RICHA H          5- K-  I5             27.00               34        28         1        17
JEZ           A BERTF          5- K-  E5             27.50               35        30         9         6
BERGER        STEVENA          5- K-  B2             27.50               32        18        15        15
BOONIN        DAVID M          5- K-  E5             27.50               35        30         0        15

DORSKIND      BRUCE K           - K-   5             27.75               33        21         2        24
COOK          ALAN  R          5- K-  J5             28.75               38        37         0         5
CARROLL       RICHARA          5- K-  D5             30.75               35        17        26         2
LEGGETT       LAWRENE          5-  -  J5             31.50               36        18         9        17
DEANS         JAMIE R          5- K-  H2             32.25               38        23         2        17

GREENBAUM     KEITH R          5- K-  D5             32.50               36        14         0        30
HALL          KEITH R          5- K-  H5             32.75               39        25         4        12
GLICKMAN      RONALDD          5- K-  G5             33.25               38        19         1        22
CALUSINE      GARY             5- K-  B              33.50               40        26         9         5
BRINN         STUART           5- K-  B5             33.75               38        17         1        24

BECHARA       DENNIS            - K-  L              34.75               40        21         4        15
EPSTEIN       JAMES A          5- K-  C5             35.00               38        12         0        30
MARITZ        JAMES A          5-  -  B              35.25               39        15         7        19
BERSHAD       RUSSELB           - K-  I5             37.75               43        21        11         5
DEPAOLO       PAUL  J          5- K-  D5             38.00               41        12         5        22

BUTLER        KENT  S          5- K-  A5             38.25               45        27         0         8
COHEN         RICHARM          5- K-  B5             38.25               42        15        18         5
BASS          ARTHUR           5-  -  E5             38.50               44        22         9         5
BIALLAS       MIKE  D          5- K-  H5             39.75               45        21         2        12
GOODMAN       ROBERTA          5- K-  J5             39.75               43        13         0        24

COSTELLO      TOM              5- K-  B5             39.75               47        29         0         4
CURRAN        JOEL  P          5- K-  K5             39.75               46        25         2         7
BRILL         JOEL  D           - C-  G5             41.00               47        24         8         1
GERSHON       STEVE A          5- K-  A5             41.00               47        24         3         6
HALLAS        RICHARD          5- K-  B5             41.00               45        16        18         1

CHAPMAN       TERRY G          5- K-  A5             41.00               46        20         0        14
BRADEORD      CHARLEC          5- K-  H              41.00               46        20        14         0
COLCORD III   EUGENEL          5- K-  C5             41.25               45        15        10        10
BELGER        LARRY A          5- K-  G5             41.50               46        18         1        15
FULOP         RONALDG          5- K-  E5             43.25               48        19        10         3

        .                                              .                                   .
        .                                              .                                   .
        .                                              .                                   .


 HALL         ROBERTL          5- K-  D5             55.00               57         8         0        15
 CHIRLS       RICHAR           5- K-  K5             55.50               58        10        12         0
 ALTMAN       JEROLDD          5- K-  J5             56.00               58         8         2        12
 COURTNEY     WILLIAL          5- K-  I5             56.00               59        12         0         9
 LANGE        STEVE E          5-  -  D5             57.50               59         6         0        15

 CAMPBELL     FRANK D          5- K-  F5             58.00               59         4         0        17
 COX          GREGORA          5- K-  K5             58.75               61         9         0        10
 MAULE        JAMES E          5-  -  K5             61.25               65        15         0         0
 MARKS        MICHAEH          5- K-  K1             71.25               73         7         0         0
```

continued

000 0111(a)

```
                         T E S T   R U N   F O R   Q U I C K S C O R E
QUESTION   1
                 NON-READ    OMIT     -1-      2       3       4       5
BOTTOM HALF             0       1      21      0       1       3       8
TOP HALF                0       1      21      1       1       5       6
TOTAL                   0       2      42      1       2       8      14
                         DIFFICULTY =  0.518     RATIO INDEX = -0.027     PHI CORRELATION = -0.019
QUESTION   2
                 NON-READ    OMIT      1       2       3      -4-      5
BOTTOM HALF             0       0       4      0      15      15       0
TOP HALF                0       0       1      1       9      24       0
TOTAL                   0       0       5      1      24      39       0
                         DIFFICULTY =  0.457     RATIO INDEX =  0.361     PHI CORRELATION =  0.247
QUESTION   3
                 NON-READ    OMIT      1      -2-      3       4       5
BOTTOM HALF             0       0       2     29       1       0       2
TOP HALF                0       0       2     32       0       0       1
TOTAL                   0       0       4     61       1       0       3
                         DIFFICULTY =  0.855     RATIO INDEX =  0.064     PHI CORRELATION =  0.096
QUESTION   4
                 NON-READ    OMIT     -1-      2       3       4       5
BOTTOM HALF             0       0      26      4       0       1       3
TOP HALF                0       1      33      0       0       0       1
TOTAL                   0       1      59      4       0       1       4
                         DIFFICULTY =  0.822     RATIO INDEX =  0.192     PHI CORRELATION =  0.304
QUESTION   5
                 NON-READ    OMIT      1      -2-      3       4       5
BOTTOM HALF             0       1      10     18       5       0       0
TOP HALF                0       0       6     28       1       0       0
TOTAL                   0       1      16     46       6       0       0
                         DIFFICULTY =  0.587     RATIO INDEX =  0.334     PHI CORRELATION =  0.272

                                      . . .
```

continued

```
                              T E S T   R U N   F O R    Q U I C K S C O R E

QUESTION  76
                  NON-READ    OMIT     -1-      2       3       4       5
BOTTOM HALF             29       0       4      1       0       0       0
TOP HALF                20       1      10      2       1       0       1
TOTAL                   49       1      14      3       1       0       1
                              DIFFICULTY =  0.185     RATIO INDEX =  0.292     PHI CORRELATION = -0.086
QUESTION  77
                  NON-READ    OMIT      1      -2-      3       4       5
BOTTOM HALF             30       0       1      2       1       0       0
TOP HALF                21       0       2     12       0       0       0
TOTAL                   51       0       3     14       1       0       0
                              DIFFICULTY =  0.188     RATIO INDEX =  0.492     PHI CORRELATION =  0.357
QUESTION  78
                  NON-READ    OMIT      1       2       3      -4-      5
BOTTOM HALF             30       0       0      0       2       2       0
TOP HALF                23       0       2      1       2       7       0
TOTAL                   53       0       2      1       4       9       0
                              DIFFICULTY =  0.105     RATIO INDEX =  0.292     PHI CORRELATION =  0.073
QUESTION  79
                  NON-READ    OMIT      1       2       3       4      -5-
BOTTOM HALF             31       0       0      1       0       0       2
TOP HALF                23       0       0      4       1       0       7
TOTAL                   54       0       0      5       1       0       9
                              DIFFICULTY =  0.109     RATIO INDEX =  0.274     PHI CORRELATION = -0.068
QUESTION  80
                  NON-READ    OMIT      1       2       3      -4-      5
BOTTOM HALF             33       0       0      0       0       1       0
TOP HALF                23       0       0      1       0       8       3
TOTAL                   56       0       0      1       0       9       3
                              DIFFICULTY =  0.116     RATIO INDEX =  0.383     PHI CORRELATION = -0.192
```

continued

```
                         T E S T   R U N   F O R    Q U I C K S C O R E

                 D I S T R I B U T I O N   O F   S C O R E S
SCORE FREQ.
    1 (   0)
    2 (   0)
    3 (   0)
        .
        .
        .
   24 (   0)
   25 (   1) *
   26 (   0)
   27 (   1) *
   28 (   4) ****
   29 (   1) *
   30 (   0)
   31 (   1) *
   32 (   2) **
   33 (   3) ***
   34 (   2) **
   35 (   3) ***
   36 (   0)
   37 (   0)
   38 (   4) ****
   39 (   1) *
   40 (   4) ****
   41 (   6) ******
   42 (   1) *
   43 (   1) *
   44 (   1) *
   45 (   2) **
   46 (   0)
   47 (   1) *
   48 (   2) **
   49 (   4) ****
   50 (   2) **
   51 (   6) ******
   52 (   2) **
   53 (   4) ****
   54 (   1) *
   55 (   1) *
   56 (   3) ***
   57 (   0)
   58 (   2) **
   59 (   1) *
   60 (   0)
   61 (   1) *
   62 (   0)
   63 (   0)
   64 (   0)
   65 (   0)
   66 (   0)
   67 (   0)
   68 (   0)
   69 (   0)
   70 (   0)
   71 (   1) *
   72 (   0)
   73 (   0)
   74 (   0)
   75 (   0)
   76 (   0)
   77 (   0)
   78 (   0)
   79 (   0)
   80 (   0)
```

continued

```
                                   T E S T   R U N   F O R   Q U I C K S C O R E

      T E S T   S T A T I S T I C S

          NUMBER OF QUESTIONS ON TEST =      80     NUMBER OF CHOICES PER ITEM   =     5     NUMBER OF EXAMINEES =     69

          MEAN SCORE ON TEST          =   43.65     STANDARD DEVIATION OF SCORES =  9.91     VARIANCE OF SCORES  =  98.11

          ANSWER KEY =  03101313310440240124342432413012342120014240334212422141131431200330230114201343

                        CODE:
                             0 - FIRST CHOICE
                             1 - SECOND CHOICE
                             2 - THIRD CHOICE
                                 ETC.
```

SAMPLE OUTPUT—ITEMSTEP

```
                    +---------------------------------------------------------------------------+
                    |                                                                           |
                    |            T E S T      R U N      F O R      I T E M S T E P             |
                    |                                                                           |
                    +---------------------------------------------------------------------------+


     AT EACH PASS OF THIS PROGRAM THE FOLLOWING IS REPORTED:


STUDENTS' NAMES, SOCIAL SECURITY NUMBERS, AND SCORES,          HISTOGRAM OF DISTRIBUTION OF SCORES
  INCLUDING  NUMBER RIGHT                                                ALL SCORES BELOW ONE ARE TREATED AS SCORES OF ONE
             NUMBER WRONG                                                ALL SCORES ARE ROUNDED TO THE NEAREST INTEGER.
             NUMBER OMITTED
             NUMBER NOT READ                                   SUMMARY OF MAJOR TEST STATISTICS, INCLUDING
             FORMULA SCORE (CORRECTED FOR GUESSING)                      MEAN OF SCORES
                                                                         VARIANCE OF SCORES
ABRIDGED ITEM ANALYSIS FOR EACH QUESTION                                 RELIABILITY INDEX BASED ON THE COVARIANCE MATRIX
             SEE EXPLANATION OF INDICES BELOW


     ITEMSTEP NOW REMOVES QUESTIONS AS REQUESTED BY THE USER. THE TEST IS RESCORED USING ONLY THE REMAINING QUESTIONS AND A NEW
LISTING OF THE RESULTS IS GIVEN.  THIS PROCEDURE IS FOLLOWED UNTIL THE EXECUTION OF THE USER'S COMMANDS HAS BEEN COMPLETED.


MEANING OF INDICES                                            EP     - PROPORTION OF THOSE RESPONDING TO THE QUESTION, ADJUSTED
                                                                       FOR GUESSING AMONG THE EFFECTIVE NUMBER OF CHOICES
P      - THE PROPORTION OF THOSE RESPONDING TO THE QUESTION
         WHO SELECTED THE RIGHT ANSWER                        DDIFF - DAVIS DIFFICULTY INDEX, COMPUTED FROM P ADJ

Q      - THE PROPORTION OF THOSE RESPONDING TO THE QUESTION   DDISC - DAVIS DISCRIMINATION INDEX, COMPUTED FROM A WEIGHTED
         WHO SELECTED A WRONG ANSWER                                  COMBINATION OF R LAT AND COREL.  THE WEIGHTS DEPEND ON
                                                                      P ADJ.  FOR HIGH P ADJ, R LAT PREDOMINATES.  FOR P ADJ
P ADJ  - THE ESTIMATED PROPORTION OF THOSE RESPONDING TO THE          LESS THAN 0.2, COREL PREDOMINATES.
         QUESTION WHO KNEW THE RIGHT ANSWER
                                                              DICAP - DEPARTURE FROM IDEAL CHOICE APPEAL PATTERN.  IDEALLY,
Q ADJ  - THE ESTIMATED PROPORTION OF THOSE RESPONDING TO THE          HALF OF THE EXAMINEES WHO RESPOND KNOW AND CHOOSE THE
         QUESTION WHO DID NOT KNOW THE RIGHT ANSWER                   CORRECT ANSWER AND THE REST DISTRIBUTE EQUALLY AMONG
                                                                      ALL THE AVAILABLE CHOICES.  DICAP IS THE ROOT-MEAN-
VAR    - THE VARIANCE OF THE ITEM, A MEASURE OF THE NUMBER            SQUARE DEVIATION FROM THIS PATTERN PER EXAMINEE.  HIGH
         OF DISTINCTIONS IT MAKES AMONG EXAMINEES                     VALUES MAY BE DUE TO INCORRECT KEYING OR POOR PHRASING.

V-C    - THE CONTRIBUTION OF THE ITEM TO TOTAL TEST VARIANCE  W     - A ROUGH MEASURE OF THE WORTH OF AN ITEM, BASED ON A
         CONSISTING OF ITS OWN VARIANCE PLUS HALF ITS CO-             WEIGHTED COMBINATION OF DDIFF AND DDISC, 2 TO 1
         VARIANCE WITH THE OTHER ITEMS

COREL  - THE PRODUCT MOMENT CORRELATION OF ITEM SCORE WITH
         TOTAL TEST SCORE

R LAT  - THE ESTIMATED CORRELATION OF TOTAL TEST SCORE WITH
         THE LATENT VARIABLE UNDERLYING THE RESPONSE.  WHEN
         P ADJ IS LESS THAN 0.2, THIS INDEX MAY BE UNSTABLE.

R LATX- SAME AS R LAT EXCEPT THAT SCORE ON THE GIVEN ITEM IS
         EXCLUDED FROM TOTAL TEST SCORE

FZ     - FISHER'S Z TRANSFORM OF R LAT

ENC    - EFFECTIVE NUMBER OF CHOICES.  A CHOICE SELECTED BY
         NO ONE IS EFFECTIVELY NOT A CHOICE AT ALL.
```

continued

000 0111(b)

```
                                               T E S T     R U N     F O R     I T E M S T E P

          NAME                                            SCORE                    RIGHT     WRONG       OMIT  NON-READ

BOONIN       DAVID M  5 K  E52 1                          13.75                      19        21          0         0
CHIRLS       RICHAR   5 K  K50 1                          27.75                      29         5          6         0
DARLINGTON   PAUL  G  5 K  E51 1                          26.25                      29        11          0         0
BLOOM        GARY     5 K  G51                            34.00                      35         4          1         0
DEANS        JAMIE R  5 K  H25 1                          23.75                      27        13          0         0

BASS         ARTHUR   5    E51 1                          22.75                      26        13          1         0
COOK         ALAN  R  5 K  J50 1                          16.25                      21        19          0         0
COX          SHARONL  5 K      2                          32.50                      34         6          0         0
COHEN        RICHARM  5 K  B51 1                          25.00                      27         8          5         0
ALTMAN       JEROLDD  5 K  J50 1                          34.00                      35         4          1         0

DAVIS        WILFREG  5 K  L51 1                          29.50                      31         6          3         0
CALUSINE     GARY     5 K  B   1                          15.00                      20        20          0         0
BENKIN       RICHARL    K  J51 1                          26.25                      29        11          0         0
BAGLEY       BRUCE D  5 K   51 1                          32.75                      34         5          1         0
CHUDD        REEVE E  5    H51 1                          27.75                      30         9          1         0

ASLANIAN     GREGORP  5 K                                 30.00                      32         8          0         0
BIALLAS      MIKE  D  5 K  H51 1                          23.75                      27        13          0         0
CURRAN       JOEL  P  5 K  K51 1                          20.25                      24        15          1         0
BRILL        JOEL  D    C  G51 1                          27.50                      30        10          0         0
BRADFORD     CHARLEC  5 K  H                              20.25                      24        15          1         0

CAMPBELL     THEO  C  5    D51 1                          33.75                      35         5          0         0
COSTELLO     TOM      5 K  B51 1                          22.50                      26        14          0         0
BRINN        STUART   5 K  B51 1                          24.00                      27        12          1         0
BEFFA        RICHA H  5 K  I51 1                          17.50                      22        18          0         0
DEPAOLO      PAUL  J  5 K  D51 1                          26.00                      28         8          4         0

BERGER       STEVENA  5 K  B25 1                          17.75                      21        13          6         0
CHAPMAN      TERRY G  5 K  A52 1                          26.25                      29        11          0         0
DAVIS        KEN   O  5 K  B52 1                          26.25                      29        11          0         0
COX          GREGORA  5 K  K51 1                          33.75                      35         5          0         0
COURTNEY     WILLIAL  5 K  I51 1                          32.50                      34         6          0         0

BALL OU      ROGER H  5    D51 1                          31.50                      33         6          1         0
BEAUCHAMP    GEOFFRL  5 K   51 1                          27.50                      30        10          0         0
CALDWELL     DOUGLAT  5 K  G51                            31.25                      33         7          0         0
BUTLER       KENT  S  5 K  A51 1                          21.25                      25        15          0         0
 LOMQUIST    ROBERTF  5 K  J51 1                          28.75                      31         9          0         0

BECHARA      DENNIS     K  L                              23.75                      27        13          0         0
BERSHAD      RUSSELB    K  I51 1                          25.50                      28        10          2         0
CARROLL      RICHARA  5 K  D51 1                          24.50                      27        10          3         0
COLCORD III  EUGENEL  5 K  C51                            25.00                      28        12          0         0
BELGER       LARRY A  5 K  G50 1                          21.25                      25        15          0         0

CAMPBELL     FRANK D  5 K  F51 1                          37.50                      38         2          0         0
DORSKIND     BRUCE K    K    5 1                          16.50                      21        18          1         0
HALL         ROBERTL  5 K  D51 1                          35.00                      36         4          0         0
GRELLER      HOWARD   5 K  B52 1                          28.75                      31         9          0         0
GLICKMAN     RONALDD  5 K  G51 1                          23.75                      27        13          0         0

                    .                                       .                                      .
                    .                                       .                                      .
                    .                                       .                                      .
```

continued

```
                              T E S T     R U N     F O R     I T E M S T E P

1
 P= 0.609  P ADJ= 0.518  VAR= 0.363  COREL= 0.218  R LAT  = 0.329  ENC= 3.358  DDIFF =  51.  DICAP=  0.098
 Q= 0.391  Q ADJ= 0.482  V-C= 0.721  FZ    = 0.336  R LAT-X= 0.166  EP = 0.455  DDISC =  20.  W     = 20.190

2
 P= 0.565  P ADJ= 0.457  VAR= 0.384  COREL= 0.379  R LAT  = 0.593  ENC= 2.495  DDIFF =  48.  DICAP=  0.176
 Q= 0.435  Q ADJ= 0.543  V-C= 1.291  FZ    = 0.653  R LAT-X= 0.425  EP = 0.274  DDISC =  39.  W     = 39.187

3
 P= 0.884  P ADJ= 0.855  VAR= 0.160  COREL= 0.147  R LAT  = 0.259  ENC= 3.462  DDIFF =  72.  DICAP=  0.093
 Q= 0.116  Q ADJ= 0.145  V-C= 0.324  FZ    = 0.265  R LAT-X= 0.132  EP = 0.837  DDISC =  16.  W     =  4.856

4
 P= 0.855  P ADJ= 0.822  VAR= 0.187  COREL= 0.278  R LAT  = 0.463  ENC= 3.455  DDIFF =  69.  DICAP=  0.090
 Q= 0.145  Q ADJ= 0.178  V-C= 0.661  FZ    = 0.500  R LAT-X= 0.336  EP = 0.802  DDISC =  30.  W     = 20.016

5
 P= 0.667  P ADJ= 0.587  VAR= 0.342  CORFL= 0.312  R LAT  = 0.469  ENC= 2.658  DDIFF =  55.  DICAP=  0.123
 Q= 0.333  Q ADJ= 0.413  V-C= 1.002  FZ    = 0.502  R LAT-X= 0.314  EP = 0.474  DDISC =  30.  W     = 29.656

6
 P= 0.783  P ADJ= 0.728  VAR= 0.266  COREL= 0.441  R LAT  = 0.685  ENC= 4.000  DDIFF =  63.  DICAP=  0.069
 Q= 0.217  Q ADJ= 0.272  V-C= 1.248  FZ    = 0.836  R LAT-X= 0.550  EP = 0.710  DDISC =  50.  W     = 44.135

7
 P= 0.884  P ADJ= 0.855  VAR= 0.160  COREL= 0.405  R LAT  = 0.711  ENC= 3.133  DDIFF =  72.  DICAP=  0.095
 Q= 0.116  Q ADJ= 0.145  V-C= 0.890  FZ    = 0.890  R LAT-X= 0.592  EP = 0.830  DDISC =  54.  W     = 35.634

8
 P= 0.841  P ADJ= 0.801  VAR= 0.209  COREL= 0.291  R LAT  = 0.378  ENC= 3.200  DDIFF =  68.  DICAP=  0.088
 Q= 0.159  Q ADJ= 0.199  V-C= 0.581  FZ    = 0.398  R LAT-X= 0.244  EP = 0.768  DDISC =  24.  W     = 16.069

           .                                        .
           .                                        .
           .                                        .

36
 P= 0.667  P ADJ= 0.587  VAR= 0.342  COREL= 0.486  R LAT  = 0.730  ENC= 2.204  DDIFF =  55.  DICAP=  0.156
 Q= 0.333  Q ADJ= 0.413  V-C= 1.562  FZ    = 0.914  R LAT-X= 0.586  EP = 0.402  DDISC =  55.  W     = 54.140

37
 P= 0.884  P ADJ= 0.855  VAR= 0.160  COREL= 0.358  R LAT  = 0.628  ENC= 4.556  DDIFF =  72.  DICAP=  0.090
 Q= 0.116  Q ADJ= 0.145  V-C= 0.786  FZ    = 0.738  R LAT-X= 0.507  EP = 0.851  DDISC =  44.  W     = 28.770

38
 P= 0.667  P ADJ= 0.594  VAR= 0.332  COREL= 0.542  R LAT  = 0.805  ENC= 4.774  DDIFF =  55.  DICAP=  0.038
 Q= 0.333  Q ADJ= 0.406  V-C= 1.714  FZ    = 1.092  R LAT-X= 0.668  EP = 0.590  DDISC =  66.  W     = 64.360

39
 P= 0.609  P ADJ= 0.511  VAR= 0.372  COREL= 0.438  R LAT  = 0.671  ENC= 2.255  DDIFF =  51.  DICAP=  0.181
 Q= 0.391  Q ADJ= 0.489  V-C= 1.469  FZ    = 0.788  R LAT-X= 0.513  EP = 0.297  DDISC =  47.  W     = 47.454

40
 P= 0.957  P ADJ= 0.946  VAR= 0.065  COREL= 0.170  R LAT  = 0.394  ENC= 2.800  DDIFF =  84.  DICAP=  0.112
 Q= 0.043  Q ADJ= 0.054  V-C= 0.239  FZ    = 0.417  R LAT-X= 0.288  EP = 0.932  DDISC =  25.  W     = -0.871
```

continued

```
                                T E S T    R U N    F O R    I T E M S T E P

                      D I S T R I B U T I O N   O F   S C O R E S
SCORE FREQ.

    1 (  0)
    2 (  0)
    3 (  0)
    4 (  0)
    5 (  0)
    6 (  0)
    7 (  0)
    8 (  0)
    9 (  0)
   10 (  0)
   11 (  0)
   12 (  0)
   13 (  0)
   14 (  1) *
   15 (  1) *
   16 (  1) *
   17 (  2) **
   18 (  2) **
   19 (  2) **
   20 (  2) **
   21 (  3) ***
   22 (  2) **
   23 (  3) ***
   24 (  7) *******
   25 (  4) ****
   26 (  9) *********
   27 (  0)
   28 (  6) ******
   29 (  5) *****
   30 (  3) ***
   31 (  1) *
   32 (  1) *
   33 (  4) ****
   34 (  7) *******
   35 (  2) **
   36 (  0)
   37 (  0)
   38 (  1) *
   39 (  0)
   40 (  0)
```

continued

```
                              T E S T    R U N    F O R    I T E M S T E P


 S U M M A R Y   O F   T E S T


NUMBEP OF QUESTIONS =  40  NUMBER OF CHOICES = 5  MEAN SCORE =  26.14

MFAN/# OF QUFSTIONS = 0.654  VARIANCE =  30.16  STANDARD DEVIATION/# OF QUESTIONS = 0.137

PELIABILITY TEST (COVARIANCE) = 0.661

      ANSWER KFY=   0310131331044024012434243241301234212001

                   CODE:
                        0 - FIRST CHOICE
                        1 - SECOND CHOICE
                        2 - THIRD CHOICE
                            FTC.

INSTPUCTIONS INTERPRETED AS FOLLOWS:

        DROP THF  10 QUESTIONS
        WITH THE LEAST WORTH.

                     DROPPING QUESTION  11

                     DROPPING QUESTION  14

                     DROPPING QUESTION  40

                     DROPPING QUESTION  34

                     DROPPING QUESTION  16

                     DROPPING QUESTION   3

                     DROPPING QUESTION  29

                     DROPPING QUESTION  26

                     DROPPING QUESTION  18

                     DROPPING QUESTION  31
```

continued

```
                                              T E S T     R U N     F O R     I T E M S T E P

         NAME                                                SCORE                 RIGHT      WRONG        OMIT  NON-READ


BOONIN        DAVID M   5 K   E52 1                           8.75                   13         17            0         0
CHIRLS        RICHAR    5 K   K50 1                          21.00                   22          4            4         0
DARLINGTON    PAUL   G  5 K   E51 1                          18.75                   21          9            0         0
BLOOM         GARY      5 K   G51                            26.50                   27          2            1         0
DEANS         JAMIE R   5 K   H25 1                          17.50                   20         10            0         0

BASS          ARTHUR    5     E51 1                          15.25                   18         11            1         0
COOK          ALAN   R  5 K   J50 1                          11.25                   15         15            0         0
COX           SHARONL   5 K       2                          25.00                   26          4            0         0
COHEN         RICHARM   5 K   B51 1                          17.25                   19          7            4         0
ALTMAN        JFROLDD   5 K   J50 1                          28.75                   29          1            0         0

DAVIS         WILFREG   5 K   L51 1                          23.25                   24          3            3         0
CALUSINE      GARY      5 K   B   1                           6.25                   11         19            0         0
BENKIN        RICHARL     K   J51 1                          17.50                   20         10            0         0
BAGLEY        BRUCE D   5 K    51 1                          24.00                   25          4            1         0
CHUDD         REEVE E   5     H51 1                          19.00                   21          8            1         0

ASLANIAN      GREGORP   5 K                                  22.50                   24          6            0         0
BIALLAS       MIKE   D  5 K   H51 1                          16.25                   19         11            0         0
CURPAN        JOEL   P  5 K   K51 1                          14.00                   17         12            1         0
BRILL         JOEL   D    C   G51 1                          22.50                   24          6            0         0
BPADEORD      CHARLEC   5 K   H                              11.50                   15         14            1         0

CAMPBELL      THEO   C  5     D51 1                          23.75                   25          5            0         0
COSTELLO      TOM       5 K   B51 1                          17.50                   20         10            0         0
BRINN         STUART    5 K   B51 1                          15.25                   18         11            1         0
BEFFA         RICHA H   5 K   I51 1                          11.25                   15         15            0         0
DEPAOLO       PAUL   J  5 K   D51 1                          19.75                   21          5            4         0
    .                                                          .                              .
    .                                                          .                              .
    .                                                          .                              .

HALLAS        RICHARD   5 K   B51 1                          18.75                   21          9            0         0
EPSTEIN       JAMES A   5 K   C51 1                          20.00                   22          8            0         0
ERDHEIM       ERIC      5 K   A51 1                          25.00                   26          4            0         0
GREIF         STEVEND   5 K   J24 1                          26.25                   27          3            0         0



 1
  P= 0.609   P ADJ= 0.518   VAR= 0.363   COREL= 0.217   R LAT   = 0.329   ENC= 3.358   DDIFF =  51.   DICAP=  0.098
  Q= 0.391   Q ADJ= 0.482   V-C= 0.675   FZ    = 0.335   R LAT-X= 0.154   EP = 0.455   DDISC =  20.   W     = 20.132

 2
  P= 0.565   P ADJ= 0.457   VAR= 0.384   COREL= 0.362   R LAT   = 0.566   ENC= 2.495   DDIFF =  48.   DICAP=  0.176
  Q= 0.435   Q ADJ= 0.543   V-C= 1.155   FZ    = 0.615   R LAT-X= 0.386   EP = 0.274   DDISC =  37.   W     = 36.901

 4
  P= 0.855   P ADJ= 0.822   VAR= 0.187   COREL= 0.294   R LAT   = 0.489   ENC= 3.455   DDIFF =  69.   DICAP=  0.090
  Q= 0.145   Q ADJ= 0.178   V-C= 0.655   FZ    = 0.535   R LAT-X= 0.354   EP = 0.802   DDISC =  32.   W     = 21.816

 5
  P= 0.667   P ADJ= 0.587   VAR= 0.342   COREL= 0.301   R LAT   = 0.452   ENC= 2.658   DDIFF =  55.   DICAP=  0.123
  Q= 0.333   Q ADJ= 0.413   V-C= 0.906   FZ    = 0.481   R LAT-X= 0.286   EP = 0.474   DDISC =  29.   W     = 28.394

  .                                                      .                              .
  .                                                      .                              .
  .                                                      .                              .

38
  P= 0.667   P ADJ= 0.594   VAR= 0.332   COREL= 0.529   R LAT   = 0.786   ENC= 4.774   DDIFF =  55.   DICAP=  0.038
  Q= 0.333   Q ADJ= 0.406   V-C= 1.570   FZ    = 1.043   R LAT-X= 0.639   EP = 0.590   DDISC =  63.   W     = 61.522

39
  P= 0.609   P ADJ= 0.511   VAR= 0.372   COREL= 0.484   R LAT   = 0.740   ENC= 2.255   DDIFF =  51.   DICAP=  0.181
  Q= 0.391   Q ADJ= 0.489   V-C= 1.519   FZ    = 0.918   R LAT-X= 0.575   EP = 0.297   DDISC =  55.   W     = 55.305
```

continued

```
                              T E S T     R U N     F O R     I T E M S T E P

                      D I S T R I B U T I O N   O F   S C O R E S
SCORE FREQ.

    1 (   0)
    2 (   0)
    3 (   0)
    4 (   0)
    5 (   0)
    6 (   1) *
    7 (   0)
    8 (   0)
    9 (   2) **
   10 (   1) *
   11 (   3) ***
   12 (   2) **
   13 (   0)
   14 (   5) *****
   15 (   5) *****
   16 (   3) ***
   17 (   2) **
   18 (   7) *******
   19 (   5) *****
   20 (   5) *****
   21 (   7) *******
   22 (   0)
   23 (   5) *****
   24 (   4) ****
   25 (   5) *****
   26 (   2) **
   27 (   1) *
   28 (   3) ***
   29 (   1) *
   30 (   0)


 S U M M A R Y   O F   T E S T


NUMBER OF QUESTIONS =  30  NUMBER OF CHOICES = 5  MEAN SCORE =  18.97

MEAN/# OF QUESTIONS = 0.632  VARIANCE =  26.53  STANDARD DEVIATION/# OF QUESTIONS = 0.172

RELIABILITY TEST (COVARIANCE) = 0.696

     ANSWER KEY=   030131331442024342434102321200

                   CODE:
                        0 - FIRST CHOICE
                        1 - SECOND CHOICE
                        2 - THIRD CHOICE
                            ETC.

INSTRUCTIONS INTERPRETED AS FOLLOWS:

       DO  2 PASSES AND
       DROP THE 10 QUESTIONS IN EACH STEP
       WITH THE LEAST WORTH.

                    DROPPING QUESTION  35

                    DROPPING QUESTION  19

                    DROPPING QUESTION  32

                    DROPPING QUESTION  25

                    DROPPING QUESTION  13

                    DROPPING QUESTION   1

                    DROPPING QUESTION  20

                    DROPPING QUESTION  17

                    DROPPING QUESTION  24

                    DROPPING QUESTION   4
```

continued

000 0111(b)

```
                                        T E S T     R U N     F O R     I T E M S T E P

        NAME                                        SCORE               RIGHT     WRONG      OMIT  NON-READ


BOONIN      DAVID M  5 K  E52 1                     7.50                  10        10         0         0
CHIRLS      RICHAR   5 K  K50 1                    17.00                  17         0         3         0
DARLINGTON  PAUL  G  5 K  E51 1                    12.50                  14         6         0         0
BLOOM       GARY     5 K  G51                      16.50                  17         2         1         0
DEANS       JAMIE R  5 K  H25 1                    12.50                  14         6         0         0
   .                                                 .                              .
   .                                                 .                              .
   .                                                 .                              .

HALLAS      RICHARD  5 K  B51 1                    13.75                  15         5         0         0
EPSTEIN     JAMES A  5 K  C51 1                    15.00                  16         4         0         0
ERDHEIM     ERIC     5 K  A51 1                    16.25                  17         3         0         0
GREIF       STEVEND  5 K  J24 1                    17.50                  18         2         0         0
```

```
2
 P= 0.565  P ADJ= 0.457  VAR= 0.384  COREL= 0.401  R LAT  = 0.626  ENC= 2.495  DDIFF =  48.  DICAP=  0.176
 Q= 0.435  Q ADJ= 0.543  V-C= 1.090  FZ    = 0.703  R LAT-X= 0.417  EP = 0.274  DDISC =  42.  W     = 42.171

5
 P= 0.667  P ADJ= 0.587  VAR= 0.342  COREL= 0.369  R LAT  = 0.555  ENC= 2.658  DDIFF =  55.  DICAP=  0.123
 Q= 0.333  Q ADJ= 0.413  V-C= 0.948  FZ    = 0.617  R LAT-X= 0.364  EP = 0.474  DDISC =  37.  W     = 36.516

6
 P= 0.783  P ADJ= 0.728  VAR= 0.266  COREL= 0.369  R LAT  = 0.573  ENC= 4.000  DDIFF =  63.  DICAP=  0.069
 Q= 0.217  Q ADJ= 0.272  V-C= 0.835  FZ    = 0.651  R LAT-X= 0.399  EP = 0.710  DDISC =  39.  W     = 34.030

.                                                  .                                            .
.                                                  .                                            .
.                                                  .                                            .

39
 P= 0.609  P ADJ= 0.511  VAR= 0.372  COREL= 0.489  R LAT  = 0.749  ENC= 2.255  DDIFF =  51.  DICAP=  0.181
 Q= 0.391  Q ADJ= 0.489  V-C= 1.310  FZ    = 0.937  R LAT-X= 0.555  EP = 0.297  DDISC =  56.  W     = 56.418
```

```
                         D I S T R I B U T I O N   O F   S C O R E S
SCORE FREQ.

   1 (  1) *
   2 (  0)
   3 (  1) *
   4 (  3) ***
   5 (  1) *
   6 (  1) *
   7 (  4) ****
   8 (  2) **
   9 (  5) *****
  10 (  5) *****
  11 (  3) ***
  12 (  2) **
  13 (  6) ******
  14 ( 11) ***********
  15 (  5) *****
  16 (  4) ****
  17 (  4) ****
  18 (  9) *********
  19 (  2) **
  20 (  0)
```

continued

```
                              T E S T   R U N   F O R   I T E M S T E P


 S U M M A R Y   O F   T E S T


NUMBFR OF QUESTIONS =  20  NUMBER OF CHOICES = 5  MEAN SCORE =  12.27

MEAN/# OF QUESTIONS = 0.614  VARIANCE =  19.27  STANDARD DEVIATION/# OF QUESTIONS = 0.219

RELIABILITY TEST (COVARIANCE) = 0.719

      ANSWER KEY=   31313314234241031200

                    CODE:
                         0 - FIRST CHOICE
                         1 - SECOND CHOICE
                         2 - THIRD CHOICE
                             ETC.

        DROP THE 10 QUESTIONS IN EACH STEP
        WITH THE LEAST WORTH.

                         DROPPING QUESTION  33

                         DROPPING QUESTION   8

                         DROPPING QUESTION  37

                         DROPPING QUESTION  10

                         DROPPING QUESTION  28

                         DROPPING QUESTION   6

                         DROPPING QUESTION   9

                         DROPPING QUESTION  15

                         DROPPING QUESTION  22

                         DROPPING QUESTION   5



        NAME                                        SCORE           RIGHT      WRONG      OMIT  NON-READ

BOONIN       DAVID M   5 K   E52 1                   5.00               6          4         0         0
CHIRLS       RICHAR    5 K   K50 1                   9.00               9          0         1         0
DARLINGTON   PAUL   G  5 K   E51 1                   5.00               6          4         0         0
BLOOM        GARY      5 K   G51                     7.75               8          1         1         0
DEANS        JAMIE R   5 K   H25 1                   7.50               8          2         0         0
   .                                                  .                            .
   .                                                  .                            .
   .                                                  .                            .

HALLAS       RICHARD   5 K   B51 1                   6.25               7          3         0         0
EPSTEIN      JAMES A   5 K   C51 1                   8.75               9          1         0         0
ERDHFIM      ERIC      5 K   A51 1                  10.00              10          0         0         0
GREIF        STEVEND   5 K   J24 1                   8.75               9          1         0         0
```

continued

```
                                    T E S T     R U N     F O R     I T E M S T E P

 2
  P= 0.565   P ADJ= 0.457   VAR= 0.384   COREL= 0.421   R LAT   = 0.657   ENC= 2.495   DDIFF =  48.   DICAP=  0.176
  Q= 0.435   Q ADJ= 0.543   V-C= 0.730   FZ    = 0.752   R LAT-X= 0.327   EP = 0.274   DDISC =  45.   W     = 45.133

 7
  P= 0.884   P ADJ= 0.855   VAR= 0.160   COREL= 0.468   R LAT   = 0.823   ENC= 3.133   DDIFF =  72.   DICAP=  0.095
  Q= 0.116   Q ADJ= 0.145   V-C= 0.525   FZ    = 1.165   R LAT-X= 0.592   EP = 0.830   DDISC =  70.   W     = 46.341

12
  P= 0.797   P ADJ= 0.750   VAR= 0.246   COREL= 0.508   R LAT   = 0.794   ENC= 4.596   DDIFF =  64.   DICAP=  0.068
  Q= 0.203   Q ADJ= 0.250   V-C= 0.707   FZ    = 1.079   R LAT-X= 0.542   EP = 0.745   DDISC =  65.   W     = 54.905

21
  P= 0.696   P ADJ= 0.627   VAR= 0.320   COREL= 0.450   R LAT   = 0.672   ENC= 3.188   DDIFF =  57.   DICAP=  0.088
  Q= 0.304   Q ADJ= 0.373   V-C= 0.713   FZ    = 0.807   R LAT-X= 0.389   EP = 0.570   DDISC =  49.   W     = 46.855

23
  P= 0.783   P ADJ= 0.728   VAR= 0.266   COREL= 0.533   R LAT   = 0.829   ENC= 3.103   DDIFF =  63.   DICAP=  0.086
  Q= 0.217   Q ADJ= 0.272   V-C= 0.770   FZ    = 1.180   R LAT-X= 0.572   EP = 0.679   DDISC =  71.   W     = 61.402

27
  P= 0.565   P ADJ= 0.464   VAR= 0.375   COREL= 0.475   R LAT   = 0.733   ENC= 3.292   DDIFF =  48.   DICAP=  0.113
  Q= 0.435   Q ADJ= 0.536   V-C= 0.816   FZ    = 0.891   R LAT-X= 0.418   EP = 0.388   DDISC =  54.   W     = 53.490

30
  P= 0.652   P ADJ= 0.612   VAR= 0.287   COREL= 0.499   R LAT   = 0.698   ENC= 4.457   DDIFF =  56.   DICAP=  0.068
  Q= 0.348   Q ADJ= 0.388   V-C= 0.749   FZ    = 0.855   R LAT-X= 0.453   EP = 0.606   DDISC =  52.   W     = 50.063

36
  P= 0.667   P ADJ= 0.587   VAR= 0.342   COREL= 0.612   R LAT   = 0.920   ENC= 2.204   DDIFF =  55.   DICAP=  0.156
  Q= 0.333   Q ADJ= 0.413   V-C= 1.003   FZ    = 1.531   R LAT-X= 0.649   EP = 0.402   DDISC =  92.   W     = 87.927

38
  P= 0.667   P ADJ= 0.594   VAR= 0.332   COREL= 0.602   R LAT   = 0.894   ENC= 4.774   DDIFF =  55.   DICAP=  0.038
  Q= 0.333   Q ADJ= 0.406   V-C= 0.971   FZ    = 1.403   R LAT-X= 0.629   EP = 0.590   DDISC =  85.   W     = 81.552

39
  P= 0.609   P ADJ= 0.511   VAR= 0.372   COREL= 0.504   R LAT   = 0.770   ENC= 2.255   DDIFF =  51.   DICAP=  0.181
  Q= 0.391   Q ADJ= 0.489   V-C= 0.860   FZ    = 0.985   R LAT-X= 0.463   EP = 0.297   DDISC =  59.   W     = 59.292


                         D I S T R I B U T I O N   O F   S ¯ O R E S
SCORE FREQ.

     1 (   3) ***
     2 (   2) **
     3 (   d) ********
     4 (   7) *******
     5 (   8) ********
     6 (   6) ******
     7 (   2) **
     8 (  10) **********
     9 (  17) *****************
    10 (   6) ******


 S U M M A R Y   O F   T E S T


NUMBER OF QUESTIONS =  10  NUMBER OF CHOICES = 5  MEAN SCORE =   6.18

MEAN/# OF QUESTIONS = 0.618  VARIANCE =   7.84  STANDARD DEVIATION/# OF QUESTIONS = 0.280

RELIABILITY TEST (COVARIANCE) = 0.674

      ANSWER KEY=   3143240100

                    CODE:
                         0 - FIRST CHOICE
                         1 - SECOND CHOICE
                         2 - THIRD CHOICE
                             ETC.

E N D   O F   J O B
```

SAMPLE OUTPUT—ITEMRS

```
+----------------------------------------------------------------+
|                                                                |
|         T E S T   R U N   F O R   I T E M   A N A L Y S I S    |
|                                                                |
+----------------------------------------------------------------+

MEANING OF INDICES

P      - THE PROPORTION OF THOSE RESPONDING TO THE QUESTION
         WHO SELECTED THE RIGHT ANSWER

Q      - THE PROPORTION OF THOSE RESPONDING TO THE QUESTION
         WHO SELECTED A WRONG ANSWER

P ADJ  - THE ESTIMATED PROPORTION OF THOSE RESPONDING TO THE
         QUESTION WHO KNEW THE RIGHT ANSWER

Q ADJ  - THE ESTIMATED PROPORTION OF THOSE RESPONDING TO THE
         QUESTION WHO DID NOT KNOW THE RIGHT ANSWER

VAR    - THE VARIANCE OF THE ITEM, A MEASURE OF THE NUMBER
         OF DISTINCTIONS IT MAKES AMONG EXAMINEES

V-C    - THE CONTRIBUTION OF THE ITEM TO TOTAL TEST VARIANCE
         CONSISTING OF ITS OWN VARIANCE PLUS HALF ITS CO-
         VARIANCE WITH THE OTHER ITEMS

COREL  - THE PRODUCT MOMENT CORRELATION OF ITEM SCORE WITH
         TOTAL TEST SCORE

R LAT  - THE ESTIMATED CORRELATION OF TOTAL TEST SCORE WITH
         THE LATENT VARIABLE UNDERLYING THE RESPONSE.  WHEN
         P ADJ IS LESS THAN 0.2, THIS INDEX MAY BE UNSTABLE.

R LATX- SAME AS R LAT EXCEPT THAT SCORE ON THE GIVEN ITEM IS
         EXCLUDED FROM TOTAL TEST SCORE

FZ     - FISHER'S Z TRANSFORM OF R LAT

ENC    - EFFECTIVE NUMBER OF CHOICES.  A CHOICE SELECTED BY
         NO ONE IS EFFECTIVELY NOT A CHOICE AT ALL.

EP     - PROPORTION OF THOSE RESPONDING TO THE QUESTION, ADJUSTED
         FOR GUESSING AMONG THE EFFECTIVE NUMBER OF CHOICES

DDIFF  - DAVIS DIFFICULTY INDEX, COMPUTED FROM P ADJ

DDISC  - DAVIS DISCRIMINATION INDEX, COMPUTED FROM A WEIGHTED
         COMBINATION OF R LAT AND COREL.  THE WEIGHTS DEPEND ON
         P ADJ.  FOR HIGH P ADJ, R LAT PREDOMINATES.  FOR P ADJ
         LESS THAN 0.2, COREL PREDOMINATES.

DICAP  - DEPARTURE FROM IDEAL CHOICE APPEAL PATTERN.  IDEALLY,
         HALF OF THE EXAMINEES WHO RESPOND KNOW AND CHOOSE THE
         CORRECT ANSWER AND THE REST DISTRIBUTE EQUALLY AMONG
         ALL THE AVAILABLE CHOICES.  DICAP IS THE ROOT-MEAN-
         SQUARE DEVIATION FROM THIS PATTERN PER EXAMINEE.  HIGH
         VALUES MAY BE DUE TO INCORRECT KEYING OR POOR PHRASING.

W      - A ROUGH MEASURE OF THE WORTH OF AN ITEM, BASED ON A
         WEIGHTED COMBINATION OF DDIFF AND DDISC, 2 TO 1

THIS OUTPUT WAS PRODUCED BY VERSION-2 OF PROGRAM  ITEMRS, WHICH
IS THE PROPERTY OF WHARTON COMPUTATIONAL SERVICES AND IS NOT
RELEASED FOR GENERAL USE.  TO ARRANGE FOR THE PROCESSING OF
DATA BY THIS PROGRAM, CONTACT WHARTON COMPUTATIONAL SERVICES,
EXTENSION 6422.

ITEMRS WAS WRITTEN BY DANIEL ASHLER AND DANIEL BRICKLIN IN JUNE,
1969, AND WAS CHECKED IN PART BY DANIEL JAMISON AND AVRAM
REISNER.  DR. FREDERICK B. DAVIS FURNISHED INDISPENSABLE
GUIDANCE, BUT ANY ERRORS ARE THE RESPONSIBILITY OF THE AUTHORS.

ADDITIONAL INFORMATION CAN BE FOUND IN THE FINAL REPORT OF THE
COMPUTERIZED ITEM ANALYSIS PROJECT, DATED 30 JUNE 1969, AVAIL-
ABLE FROM WHARTON COMPUTATIONAL SERVICES.
```

continued

```
                                   FI°ST °
ROW  1
 31.7

ROW  2
  7.9  19.4

ROW  3
 23.9  11.0  36.4

ROW  4
 23.6  12.2  31.4  37.3

ROW  5
 17.6   4.7  16.3  17.6  26.9

ROW  6
 23.6  13.8  25.1  26.4  14.4  32.6

ROW  7
 23.2  10.4  32.6  32.3  20.4  26.1  38.3

ROW  8
 25.4  11.0  30.1  31.4  19.4  25.1  32.6  36.4

ROW  9
 20.4   8.7  26.6  26.3  15.9  24.7  29.1  26.6  33.5

ROW 10
 19.8  10.1  22.9  21.1  13.8  21.1  23.9  21.4  20.7  29.8

ROW 11
 22.9  13.2  27.6  28.9  13.8  27.3  30.1  27.6  24.1  26.7  34.5

ROW 12
 22.0  13.8  25.1  27.9  14.4  23.2  24.5  25.1  21.6  17.9  25.7  32.6

ROW 13
 28.3  12.2  34.5  35.7  22.3  29.5  37.0  34.5  31.0  25.7  32.0  29.5  42.0

ROW 14
 14.8   3.2  14.8  16.1  10.1  14.5  17.3  16.4  15.7  13.9  17.0  11.4  20.8  26.1

ROW 15
  9.4   0.9  10.9  12.2   4.6   7.5  10.3   9.4  11.6   5.0   6.9  10.6  13.7   7.5  20.3

ROW 16
 15.1   5.1  18.2  19.5  11.9  14.8  19.2  18.2  17.9  15.7  17.3  13.2  19.5  10.7   5.3  25.1

ROW 17
 27.0  12.6  28.6  29.8  17.9  28.3  29.5  30.1  25.1  21.4  29.2  28.3  34.5  17.9  10.6  13.6  36.4

ROW 18
 28.9  11.3  33.6  33.3  19.7  28.6  34.5  33.6  28.5  23.2  29.5  28.6  39.5  19.8  12.8  17.0  33.6  40.1

ROW 19
 22.6  11.3  28.9  28.6  19.7  23.9  31.4  32.0  23.8  20.1  24.8  23.9  33.3  15.1   8.1  17.0  28.9  32.3  35.4

ROW 20
 19.8  11.6  22.9  24.2  13.8  19.5  23.9  22.9  16.3  14.2  20.4  21.1  25.7  10.7   5.3  12.6  24.5  24.8  24.8  29.8
                    .                                   .                                   .
                    .                                   .                                   .
                    .                                   .                                   .
ROW 40
 23.5  13.4  31.3  32.6  19.1  26.3  35.4  32.9  27.8  22.6  28.8  26.3  37.2  17.3  10.2  17.9  29.8  34.8  31.6  22.6
 12.5  25.3  26.6  30.6   8.9  25.6  22.8  19.4  29.8  25.1  28.1  31.5  19.4   6.9  20.6  11.6  33.8  18.1  21.3  38.2
```

continued

```
                              COV-MATX

ROW  1
 0.30

ROW  2
-0.01  0.38

ROW  3
 0.02  0.02  0.19

ROW  4
-0.01  0.04  0.05  0.16

ROW  5
 0.06 -0.04 -0.05 -0.03  0.36

ROW  6
 0.09  0.12  0.02  0.03 -0.03  0.28

ROW  7
-0.04 -0.01  0.05  0.02  0.02  0.00  0.13

ROW  8
 0.05  0.02  0.04  0.05  0.03  0.02  0.05  0.19

ROW  9
-0.01 -0.01  0.03 -0.00 -0.01  0.07  0.04  0.03  0.25

ROW 10
 0.06  0.06  0.04 -0.02  0.00  0.07  0.03  0.01  0.04  0.33

ROW 11
 0.03  0.09  0.03  0.04 -0.07  0.12  0.05  0.03  0.01  0.17  0.24

ROW 12
 0.05  0.12  0.02  0.07 -0.03  0.06 -0.04  0.02 -0.00 -0.00  0.08  0.28

ROW 13
-0.0  -0.0  -0.0  -0.0  -0.0  -0.0  -0.0  -0.0  -0.0  -0.0  -0.0  -0.0  -0.0

ROW 14
 0.02 -0.07 -0.05 -0.04 -0.02 -0.00 -0.02 -0.02  0.01  0.03  0.03 -0.08 -0.0   0.38

ROW 15
 0.00 -0.07 -0.01  0.01 -0.06 -0.05 -0.04 -0.05  0.03 -0.08 -0.09  0.02 -0.0   0.02  0.38

ROW 16
 0.05 -0.01  0.05  0.07  0.04  0.03  0.05  0.05  0.08  0.09  0.06 -0.01 -0.0   0.03 -0.03  0.38

ROW 17
 0.09  0.06  0.01  0.01 -0.01  0.10 -0.02  0.04 -0.01  0.01  0.07  0.10 -0.0   0.02 -0.02 -0.06  0.19

ROW 18
 0.05 -0.00  0.03 -0.01 -0.03  0.02 -0.01  0.03 -0.02 -0.02 -0.01  0.02 -0.0   0.01 -0.00 -0.03  0.03  0.07

ROW 19
 0.01  0.04  0.04  0.01  0.05  0.01  0.05  0.11 -0.02 -0.01 -0.01  0.01 -0.0  -0.03 -0.07  0.04  0.04  0.02  0.22

ROW 20
 0.06  0.10  0.04  0.05  0.00  0.03  0.03  0.04 -0.06 -0.04  0.02  0.07 -0.0  -0.05 -0.07  0.02  0.08  0.01  0.11  0.33
                    .                                         .                                   .
                    .                                         .                                   .
                    .                                         .                                   .
ROW 40
-0.04  0.06  0.02  0.02 -0.02  0.00  0.06  0.06  0.01 -0.01  0.01  0.00 -0.0  -0.03 -0.05  0.01 -0.02 -0.01  0.05 -0.01
-0.07  0.00  0.04  0.05  0.02  0.04  0.02  0.08 -0.02  0.05  0.01  0.01  0.05  0.04  0.02  0.02  0.05  0.03 -0.01  0.11
```

continued

COR-MATX

```
ROW  1
1.000

ROW  2
-.026 1.000

ROW  3
0.066 0.083 1.000

ROW  4
-.052 0.174 0.270 1.000

ROW  5
0.187 -.116 -.178 -.135 1.000

ROW  6
0.303 0.379 0.091 0.140 -.089 1.000

ROW  7
-.193 -.044 0.331 0.131 0.084 0.009 1.000

ROW  8
0.221 0.083 0.222 0.270 0.105 0.091 0.331 1.000

ROW  9
-.041 -.022 0.126 -.009 -.038 0.265 0.234 0.126 1.000

ROW 10
0.187 0.171 0.168 -.087 0.012 0.230 0.134 0.021 0.138 1.000

ROW 11
0.125 0.304 0.149 0.196 -.254 0.441 0.256 0.149 0.049 0.593 1.000

ROW 12
0.176 0.379 0.091 0.312 -.089 0.213 -.181 0.091 -.013 -.012 0.298 1.000

ROW 13
0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   0.0   1.000

ROW 14
0.060 -.181 -.198 -.153 -.060 -.005 -.102 -.059 0.029 0.077 0.094 -.233 0.0   1.000

ROW 15
0.009 -.198 -.032 0.046 -.172 -.157 -.190 -.170 0.109 -.230 -.285 0.071 0.0   0.045 1.000

ROW 16
0.141 -.039 0.196 0.276 0.103 0.081 0.211 0.196 0.267 0.253 0.193 -.032 0.0   0.070 -.067 1.000

ROW 17
0.376 0.221 0.028 0.060 -.036 0.411 -.132 0.222 -.043 0.021 0.322 0.411 0.0   0.079 -.059 -.216 1.000

ROW 18
0.375 -.030 0.228 -.082 -.175 0.138 -.073 0.228 -.116 -.150 -.108 0.138 0.0   0.043 -.017 -.194 0.228 1.000

ROW 19
0.024 0.134 0.183 0.033 0.182 0.050 0.290 0.548 -.074 -.023 -.054 0.050 0.0   -.108 -.230 0.129 0.183 0.200 1.000

ROW 20
0.187 0.276 0.168 0.231 0.012 0.109 0.134 0.168 -.221 -.114 0.069 0.230 0.0   -.132 -.209 0.045 0.315 0.092 0.392 1.000
            .                                   .                              .
            .                                   .                              .
            .                                   .                              .
ROW 40
-.212 0.314 0.122 0.158 -.084 0.021 0.523 0.390 0.047 -.036 0.066 0.021 0.0   -.141 -.238 0.070 -.145 -.080 0.343 -.036
-.339 0.005 0.214 0.316 0.086 0.196 0.111 0.392 -.145 0.302 0.053 0.084 0.227 0.238 0.111 0.119 0.362 0.176 -.053 1.000
```

000 0111(c)

continued

000 0111(c)

```
                              C L U S T E R   A N A L Y S I S

PASS NUMBER 1     LOWFST FZ = 0.572
        CLUSTER  2 =    2 (P=0.3)
        CLUSTER  3 =    3 (P=0.8)
        CLUSTER  4 =    4 (P=0.9)
        CLUSTER  5 =    5 (P=0.5)
        CLUSTER  9 =    9 (P=0.7)
        CLUSTER 11 =   11 (P=0.8)  10 (P=0.6)
        CLUSTER 12 =   12 (P=0.7)
        CLUSTER 13 =   13 (P=1.0)
        CLUSTER 14 =   14 (P=0.5)
        CLUSTER 15 =   15 (P=0.3)
        CLUSTER 16 =   16 (P=0.5)
        CLUSTER 17 =   17 (P=0.8)
        CLUSTER 18 =   18 (P=0.9)
        CLUSTER 19 =   19 (P=0.8)   8 (P=0.8)
        CLUSTER 20 =   20 (P=0.6)
        CLUSTER 21 =   21 (P=0.4)
        CLUSTER 22 =   22 (P=0.7)
        CLUSTER 23 =   23 (P=0.7)
                  .
                  .
                  .
PASS NUMBER 5     LOWEST FZ =-0.027
        CLUSTER 40 =   40 (P=0.9)   7 (P=0.9)  37 (P=0.9)  19 (P=0.8)   8 (P=0.8)  32 (P=0.8)  30 (P=0.6)  28 (P=0.4)  24 (P=0.8)

                       20 (P=0.6)  39 (P=0.6)  18 (P=0.9)  23 (P=0.7)   3 (P=0.8)  35 (P=0.5)  34 (P=0.1)  16 (P=0.5)   9 (P=0.7)
                       33 (P=0.5)  31 (P=0.7)   6 (P=0.7)  29 (P=0.8)   1 (P=0.7)  11 (P=0.8)  10 (P=0.6)  26 (P=0.6)   2 (P=0.3)
                       17 (P=0.8)  12 (P=0.7)  21 (P=0.4)  25 (P=0.2)   4 (P=0.9)  22 (P=0.7)   5 (P=0.5)  13 (P=1.0)  38 (P=0.4)
                       36 (P=0.3)  14 (P=0.5)  27 (P=0.6)  15 (P=0.3)
```

continued

```
                                          TEST RUN FOR ITEM ANALYSIS

        NAME                                      SCORE            RIGHT     WRONG      OMIT  NON-READ

ADRIANCE     JOHN  A  5 K  A51 1                  28.00               30         8         2         0
BACARDI      TOTEN D  5 K  C51                    26.25               29        11         0         0
BECKER       BARRY C  5 K  L51                    33.75               35         5         0         0
BELL        WILLIAMG  5    D50 1                  13.25               18        19         3         0
BERMAN       DENNISN  5    A51 1                  30.00               32         8         0         0
BIANCHI      DAVID A  5 K  D51 1                  25.00               28        12         0         0

BINIK        ADRIENL  5 K  C51 2                  29.50               31         6         3         0
BLOCK        ALAN  M    K  K51 1                  31.25               33         7         0         0
BRESKY       ROBERTJ  5 K  L51 1                  32.50               34         6         0         0
BUTKUS       STEPHER  5 K  F51 1                  23.25               26        11         3         0
CADGE        DONALDA  5    E51 1                  31.25               33         7         0         0
                .                                   .                                      .
                .                                   .                                      .
                .                                   .                                      .
MARKS        ROGER H  5 K  J51 1                  15.00               20        20         0         0
MILLIGAN     DAVID S  5 K  A51 1                  15.25               20        19         1         0
NEMEROVSKI   STEVENH  5 K  K51 1                  18.75               23        17         0         0
PETERSON     DAVID T  5 K  G51 1                  27.50               30        10         0         0
POLICICCHIO  JOSEPHB  5 K  E51 1                  20.00               24        16         0         0

PRUZANSKY    LEONARJ  5 K  F22 1                  16.25               21        19         0         0
REICHLE      RICHARA  5 K  A52 1                  28.75               31         9         0         0
REINER       ROBERTM  5 K  B52 1                  31.25               33         7         0         0
REKEY        ANDRASG  5 K  B51 1                  24.25               27        11         1         1
RENDINE      BARBARJ  5 K  D51 2                  25.00               28        12         0         0

REYNOLDS     ROBERTJ  5 K  B51 1                  30.00               32         8         0         0
```

```
                                          TEST RUN FOR ITEM ANALYSIS

        NAME                                      SCORE            RIGHT     WRONG      OMIT  NON-READ


BELL        WILLIAMG  5    D50 1                  13.25               18        19         3         0
MARKS        ROGER H  5 K  J51 1                  15.00               20        20         0         0
CASTELLANO   MIKE  N  5 K  B51 1                  15.25               20        19         1         0
MILLIGAN     DAVID S  5 K  A51 1                  15.25               20        19         1         0
GARST        SM UELD  5 K  D50 1                  16.25               21        19         0         0

PRUZANSKY    LEONARJ  5 K  F22 1                  16.25               21        19         0         0
DELLHEIM     GERALDL  5    D51 1                  17.50               21        14         5         0
NEMEROVSKI   STEVENH  5 K  K51 1                  18.75               23        17         0         0
POLICICCHIO  JOSEPHB  5 K  E51 1                  20.00               24        16         0         0
CAPOB                 5 K  K19 1                  20.25               24        15         1         0
                .                                   .                                      .
                .                                   .                                      .
                .                                   .                                      .
REINER       ROBERTM  5 K  B52 1                  31.25               33         7         0         0
FINKELSTEIN  LAWREN   5 K  C52 1                  31.25               33         7         0         0
CADGE        DONALDA  5    E51 1                  31.25               33         7         0         0
BLOCK        ALAN  M    K  K51 1                  31.25               33         7         0         0
BRESKY       ROBERTJ  5 K  L51 1                  32.50               34         6         0         0

BECKER       BARRY C  5 K  L51                    33.75               35         5         0         0
GREENE       BRUCE R  5 K  B52 1                  35.00               36         4         0         0
```

continued

```
                              TEST RUN FOR ITEM ANALYSIS

QUESTION   1
                   NON-READ     OMIT      -1-      2       3       4       5
    BOTTOM  20%        0          0        3       2       0       0       3
    SECOND  20%        0          0        6       0       0       1       2
    THIRD   20%        0          0        7       0       0       1       0
    FOURTH  20%        0          0        8       0       0       1       0
    TOP     20%        0          0        7       0       0       1       0
   TOTAL               0          0       31       2       0       4       5
   P= 0.738  P ADJ= 0.673  VAR= 0.302  COREL= 0.412  R LAT  = 0.628  ENC= 3.689  DDIFF =  59.  DICAP=  0.072
   Q= 0.262  Q ADJ= 0.327  V-C= 1.277  FZ    = 0.734  R LAT-X= 0.489  EP = 0.641  DDISC =  44.  W     = 41.094
                                                                                                     (GOOD)
QUESTION   2
                   NON-READ     OMIT       1       2       3      -4-      5
    BOTTOM  20%        0          0        0       0       7       1       0
    SECOND  20%        0          0        1       0       4       4       0
    THIRD   20%        0          1        0       0       5       2       0
    FOURTH  20%        0          0        0       0       5       4       0
    TOP     20%        0          0        0       0       1       7       0
   TOTAL               0          1        1       0      22      18       0
                                                                                                  (RE-WORD)
   P= 0.429  P ADJ= 0.292  VAR= 0.378  COREL= 0.418  R LAT  = 0.748  ENC= 2.091  DDIFF =  38.  DICAP=  0.291
   Q= 0.571  Q ADJ= 0.708  V-C= 1.446  FZ    = 0.801  R LAT-X= 0.566  EP =-0.073  DDISC =  48.  W     = 43.344
                                                                                                     (GOOD)
        .                         .
        .                         .
        .                         .

QUESTION  40
                   NON-READ     OMIT       1      -2-      3       4       5
    BOTTOM  20%        0          0        0       6       1       0       1
    SECOND  20%        1          0        0       8       0       0       0
    THIRD   20%        0          0        0       7       0       0       1
    FOURTH  20%        0          0        0       9       0       0       0
    TOP     20%        0          0        0       8       0       0       0
   TOTAL               1          0        0      38       1       0       2
   P= 0.927  P ADJ= 0.909  VAR= 0.106  COREL= 0.370  R LAT  = 0.732  ENC= 2.800  DDIFF =  78.  DICAP=  0.105
   Q= 0.073  Q ADJ= 0.091  V-C= 0.678  FZ    = 0.934  R LAT-X= 0.625  EP = 0.886  DDISC =  56.  W     = 28.851
                                                                                                     (FAIR)
```

continued

```
                                        T E S T   R U N   F O R   I T E M   A N A L Y S I S

                      D I S T R I B U T I O N   O F   S C O R E S
SCORE FREQ.

    1 (   0)
    2 (   0)
    3 (   0)
    4 (   0)
    5 (   0)
    6 (   0)
    7 (   0)
    8 (   0)
    9 (   0)
   10 (   0)
   11 (   0)
   12 (   0)
   13 (   1) *
   14 (   0)
   15 (   3) ***
   16 (   2) **
   17 (   0)
   18 (   1) *
   19 (   1) *
   20 (   2) **
   21 (   0)
   22 (   0)
   23 (   3) ***
   24 (   1) *
   25 (   3) ***
   26 (   4) ****
   27 (   0)
   28 (   5) *****
   29 (   4) ****
   30 (   5) *****
   31 (   4) ****
   32 (   0)
   33 (   1) *
   34 (   1) *
   35 (   1) *
   36 (   0)
   37 (   0)
   38 (   0)
   39 (   0)
   40 (   0)
```

continued

```
                         T E S T   R U N   F O R   I T E M   A N A L Y S I S

S U M M A R Y   O F   T E S T

     NUMBER OF EXAMINEES           =    42    NUMBER OF CHOICES PER ITEM =      5   NUMBER OF QUESTIONS ON TEST        =    40

     STANDARD DEVIATION OF SCORES  =  5.63    MEAN SCORE ON TEST         =  25.47   TEST RELIABILITY (K-R FORMULA 20)  = 0.675

     STANDARD ERROR OF MEAS.       =  1.35    TEST VARIANCE              =  31.74   VARIANCE ERROR OF MEAS.            = 1.830

                                                                                    TEST RELIABILITY (COVARIANCE)      = 0.671

     ANSWER KEY = 0310131331044024012434243241301234212001

     CODE:
          0 = FIRST CHOICE
          1 = SECOND CHOICE
          2 = THIRD CHOICE
              ETC.
```

COST ESTIMATE

For data submitted on answer sheets, the total cost will consist of a setup charge and a per-student charge.

	QUICKSCORE	ITEMSTEP or ITEMRS
Setup charge	$ 7.50	$ 15.00
Per-student charges		
0– 50 questions	$ .045	$ .075
51–100 questions	.075	.12
101–150 questions	.105	.165
151–200 questions	.135	.21

There will also be an extra charge for any additional work caused by crumpled or poorly marked answer sheets and/or unusual layouts requiring board rewiring, etc., at the rate of $10/hour of labor. The unusual-layout charge can be avoided by using the answer sheets available from Wharton Computational Services at $30.00/ thousand.

For data submitted on punched cards, the setup charge is eliminated and the per-student charges are reduced by $.02. There is an extra charge for the punch option of ITEMRS.

CONTENTS—QUICKSCORE/ITEMSTEP/ITEMRS

pages
1– 4 Identification & Abstract
5– 6 User Instructions–QUICKSCORE
7– 8 –ITEMSTEP
9 –ITEMRS
11 Sample Input–QUICKSTEP/ITEMSTEP/ITEMRS
12–18 Sample Output–QUICKSTEP
19–28 –ITEMSTEP
29–37 –ITEMRS
39 Cost—Contents

000 0112

DESCRIPTIVE TITLE	Page Plotter Using Line Printer
CALLING NAME	LPLOT
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	James J. Hurt, for 7044 Version Wayne Robinson, conversion for 360 University Computer Center The University of Iowa
LANGUAGE	FORTRAN IV and Assembler Language for IBM 360/65
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

The subroutine LPLOT produces on the line printer a plot of from one to eight ordinate arrays versus one abscissa array on the same plot. A plot summary of the maximums, minimums, and scale per inch for the abscissa and ordinates is printed below the plot.

USER INSTRUCTIONS

Calling Sequence

CALL LPLOT (N,X,Y1,Y2,...,YM)

where

N = the number of elements in each array. These arrays must all have the same dimension, N.

X = the name of a singly-dimensioned real (single precision) array containing the values of the abscissa.

Y1,Y2,...,YM = the names of from one to eight singly-dimensioned real (single precision) arrays containing the values of the ordinates. The ordinate value Y(i) should correspond to the abscissa value X(i) for i = 1,2,...,M. From one to eight ordinate arrays may be specified. All will be plotted versus the one abscissa array on the same plot.

Alternate Calling Sequence

CALL LPLOT ('heading',L,N,X,Y1,Y2,...,YM)

where X,Y1,Y2,...,YM are the same as above, and where

heading = any alphanumeric title information to be printed as a heading above the plot. The heading should consist of an even number, but not more than 84, characters, and it must be enclosed in apostrophes.

L = the number of characters in the heading, including all blanks but not including the apostrophes. L must be an *even* number.

Note: If the alternate calling sequence is used, the plot will automatically begin at the top of a new page. This is *not* true for the calling sequence without the heading option. If the heading is not used, the calling program should issue a WRITE statement with a carriage control of '1' before calling LPLOT to ensure that the entire plot will be printed on one page.

Method

The output of LPLOT is a plot 51 lines long and 101 columns wide with coordinate axes and a summary of the maxima, minima, and scale of the abscissa and ordinates. The entire output page

continued

b

consists of 57 lines with the heading and 56 lines without it. The plot is scaled so that the minimum ordinate is plotted at the bottom of the page, the maximum ordinate at the top, the minimum abscissa at the extreme left, and the maximum abscissa at the extreme right. The increment between rows of the plot is (1/50) (maximum value of all ordinates – minimum value of all ordinates). Likewise, the increment between columns is (1/100) (maximum of all abscissas – minimum of all abscissas).

Once the maxima, minima, and increments have been found, the arrays are plotted one line at a time. For each line, the value of each ordinate is tested, and if it falls within the range of the particular line being produced, then it is plotted in the column determined by the corresponding abscissa.

If only one ordinate array is to be plotted, all points are plotted as asterisks. If more than one is to be plotted, the points of the first are plotted with the character 1, the points of the second with 2, etc. If a point on the plot represents a point in two or more ordinate arrays, that point is plotted with an asterisk to indicate the intersection.

Restrictions

The number of values in the ordinate and abscissa arrays must be at least one and less than 2^{16}. If the heading option is used, the heading should contain an even number ($\leq$ 84) of characters. If LPLOT is called with too few or too many arguments, an error message is printed.

LPLOT does not produce graphs of extremely high accuracy, but it will provide a pictorial representation of the data in-line with the other output. The larger the differences between the maximum and minimum ordinate and the maximum and minimum abscissa, the less accurate the graph will be.

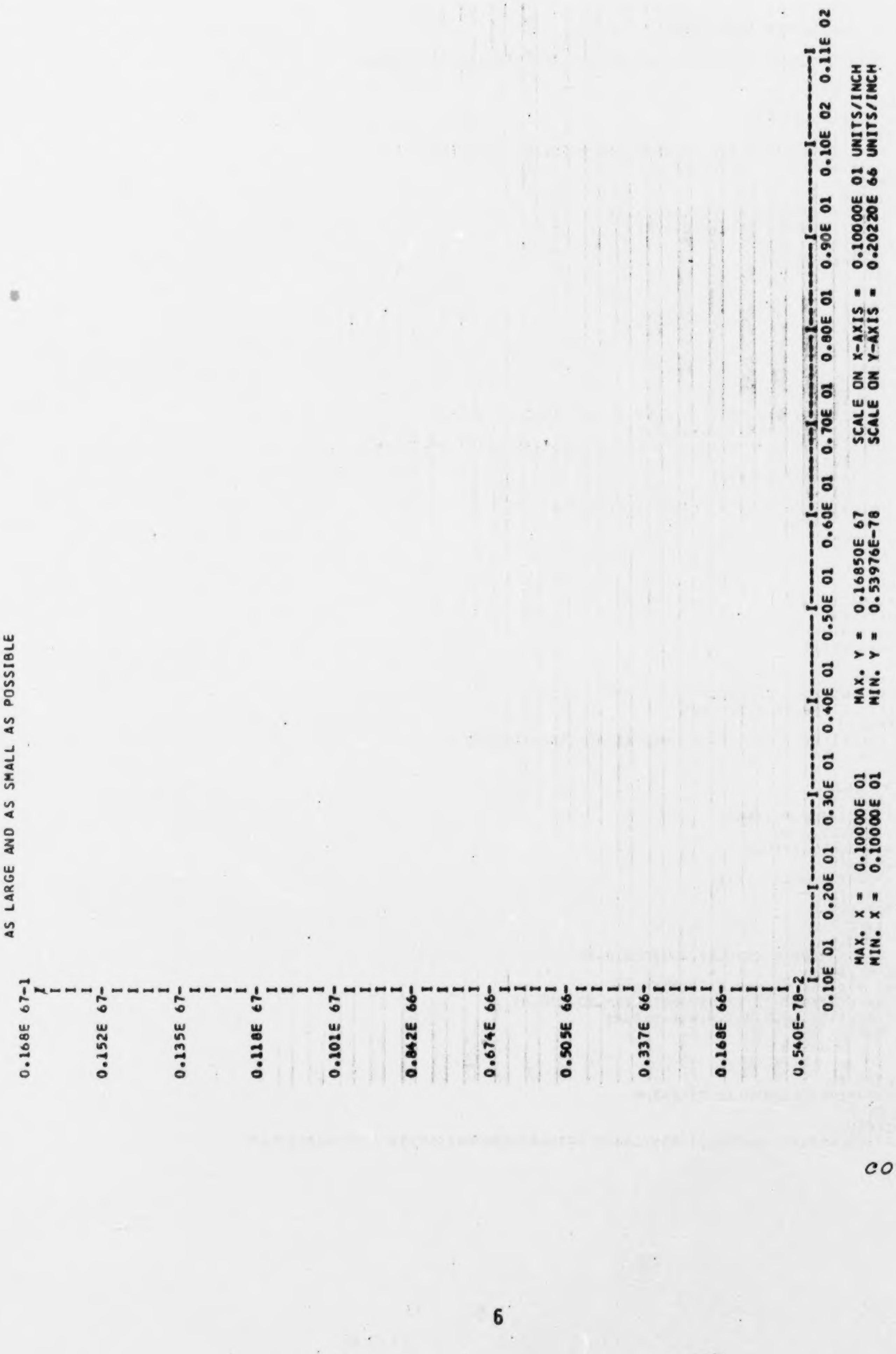
The subroutine uses about 5.5K bytes of core storage. Storage for the ordinate and abscissa arrays is not included in this figure.

LPLOT, an assembler language routine, and LV8PLT, a FORTRAN IV subroutine called by LPLOT are utilized in a call to LPLOT.

SAMPLE INPUT

```
C
C   PROGRAM TC TEST LPLOT SUBROUTINE
C
      REAL B(360),W(360),X(360),Y(360),Z(360),S(2002),R(2002),T(101),
     &V(101)
      CO 9 I=1,360
    9 B(I)=I
      READ (5,80) X(1),Y(1)
   80 FCRMAT (2Z8)
      CALL LFLCT('AS LARGE ANC AS SMALL AS POSSIBLE ',34,1,B,X,Y)
      CO 10 I=1,200
      R(I)=0.0
   10 S(I)=100.0
      CALL LFLCT('HORIZONTAL',10,200,B,R)
      CALL LFLCT('VERTICAL',8,200,S,B)
      CC 12 I=1,101
      A=(I-1.C)/100.0
      R(I)=A
      S(I)=A*A
      T(I)=A*A*A
      V(I)=A**4
      W(I)=A**5
      X(I)=A**6
      Y(I)=A**7
   12 Z(I)=A**8
      WRITE (6,13)
   13 FCRMAT('1',24X,'EXPONENTIAL CURVES FROM ZERC TO ONE')
      CALL LFLCT(101,R,R,S,T,V,W,X,Y,Z)
      CALL LFLCT('THE SAME CURVES CN AN EXPONENTIAL BASE',38,101,W,R,S,
     1T,V,W,X,Y,Z)
      READ (5,79)(R(I),I=1,20)
   79 FCRMAT (20F4.2)
      CALL LPLOT('SOME RANDCM POINTS',18,20,B,R)
      RAD=57.29577923
      CC 11 I=1,360
      W(I)=0.0
      A=I-1.0
      D=A/RAD
      X(I)=SIN(D)
      Y(I)=CCS(D)
      IF(Y(I)) 4,5,4
    4 Z(I)=X(I)/Y(I)
      GC TO 6
    5 Z(I)=2.0
    6 IF (Z(I).GT.2.0) Z(I)=2.0
      IF (Z(I).LT.-2.C) Z(I)=-2.0
   11 CONTINUE
      CALL LFLCT('SIN, CCSINE, & TANGENT',22,360,B,W,X,Y,Z)
      DO 14 I=1,1000

      R(I)=I/10.0
      A=50.0-R(I)
   14 S(I)=SCRT(2500.0-(A*A))
      DC 15 I=1001,2000
      R(I)=(I-1000.0)/10.0
      A=50.0-R(I)
   15 S(I)=-SCRT(2500.0-(A*A))
      R(2001)=-6.0
      R(2002)=107.0
      S(2001)=0.0
      S(2002)=0.0
      CALL LFLCT('SUPER CIRCLE',12,2002,R,S)
      WRITE (6,145)
  145 FCRMAT ('1TEST CF ERROR ROUTINES')
      CALL LFLCT('TEST OF ERROR ROUTINES',22,100,X)
      CALL LPLCT(100,B,R,S,T,V,W,X,Y,Z,B)
      CALL EXIT
      END
/*
//LKED.SYSLIB  DD
//     DD  DSNAME=SYS1.MATHLIB,DISP=SHR
//GO.SYSIN DD *
77FFFFFF0C1CC000
98765432768565726583572314732027353572473583472473583470027357020273573243271632
/*

TEST DECK
//
```

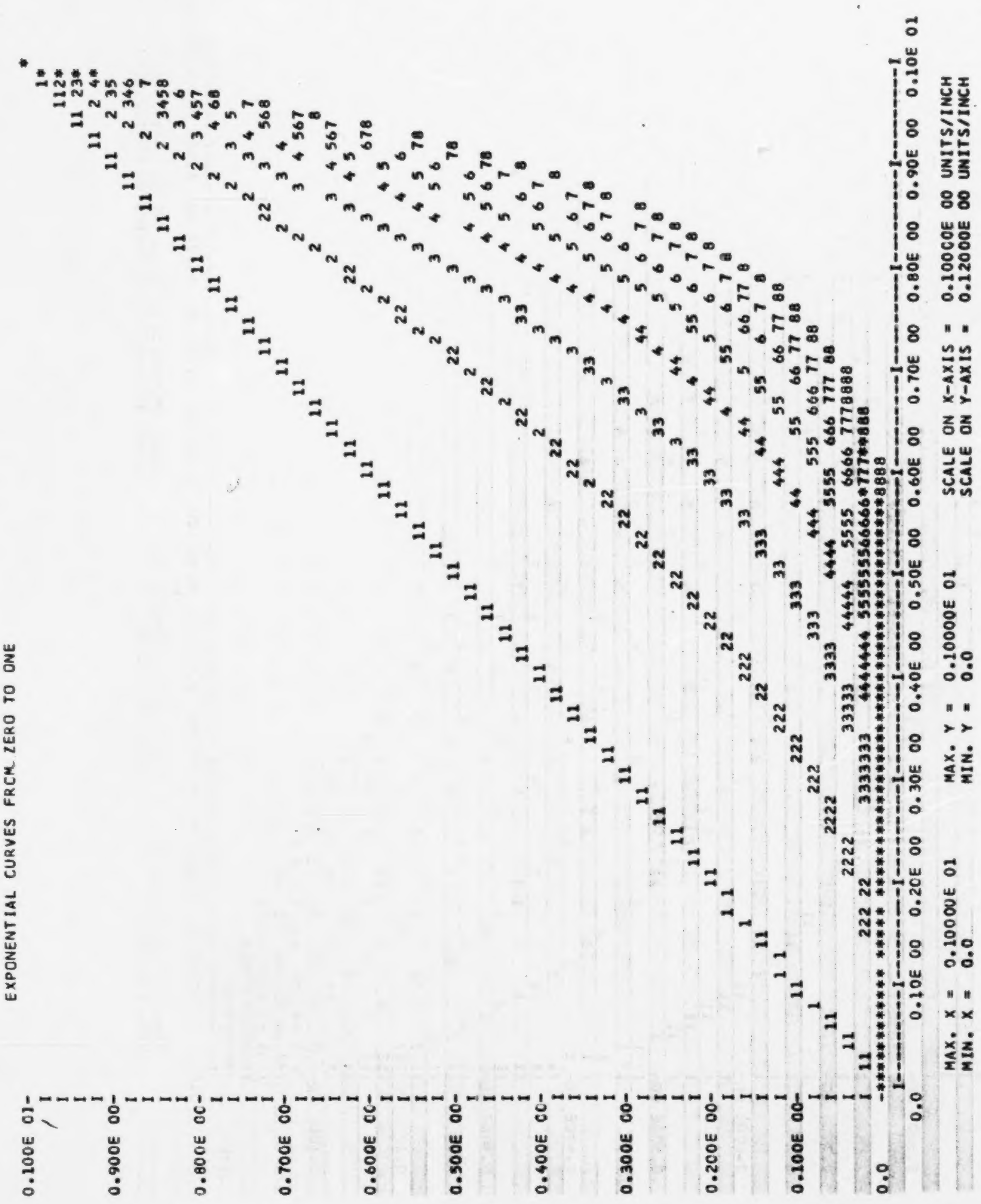
SAMPLE OUTPUT



continued

000 0112

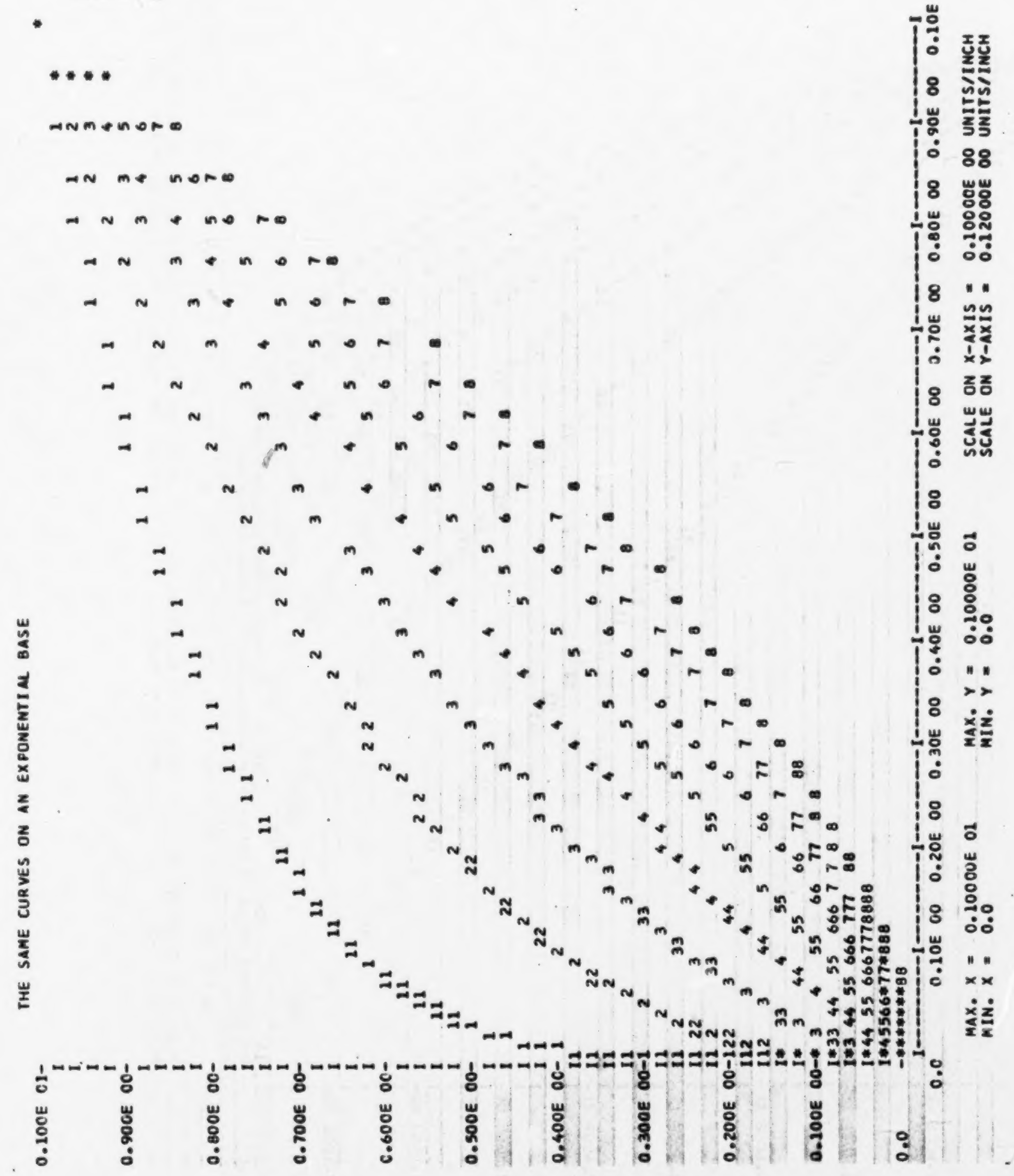
EXPONENTIAL CURVES FRCM ZERO TO ONE



```
MAX. X =  0.10000E 01     MAX. Y =  0.10000E 01     SCALE ON X-AXIS =  0.10000E 00 UNITS/INCH
MIN. X =  0.0             MIN. Y =  0.0             SCALE ON Y-AXIS =  0.12000E 00 UNITS/INCH
```

continued

000 0112

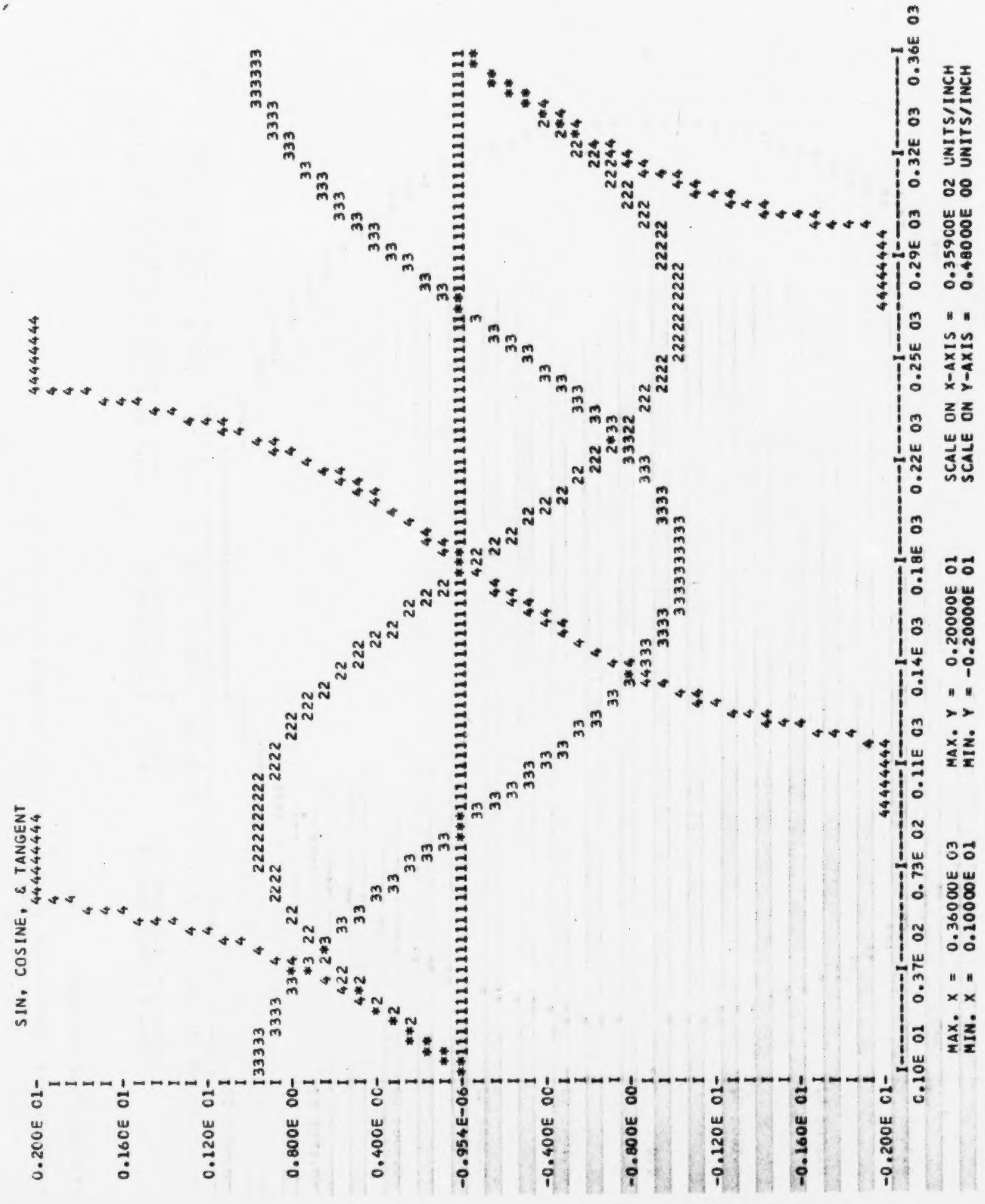


THE SAME CURVES ON AN EXPONENTIAL BASE

MAX. X = 0.10000E 01 MAX. Y = 0.10000E 01 SCALE ON X-AXIS = 0.10000E 00 UNITS/INCH
MIN. X = 0.0 MIN. Y = 0.0 SCALE ON Y-AXIS = 0.12000E 00 UNITS/INCH

continued

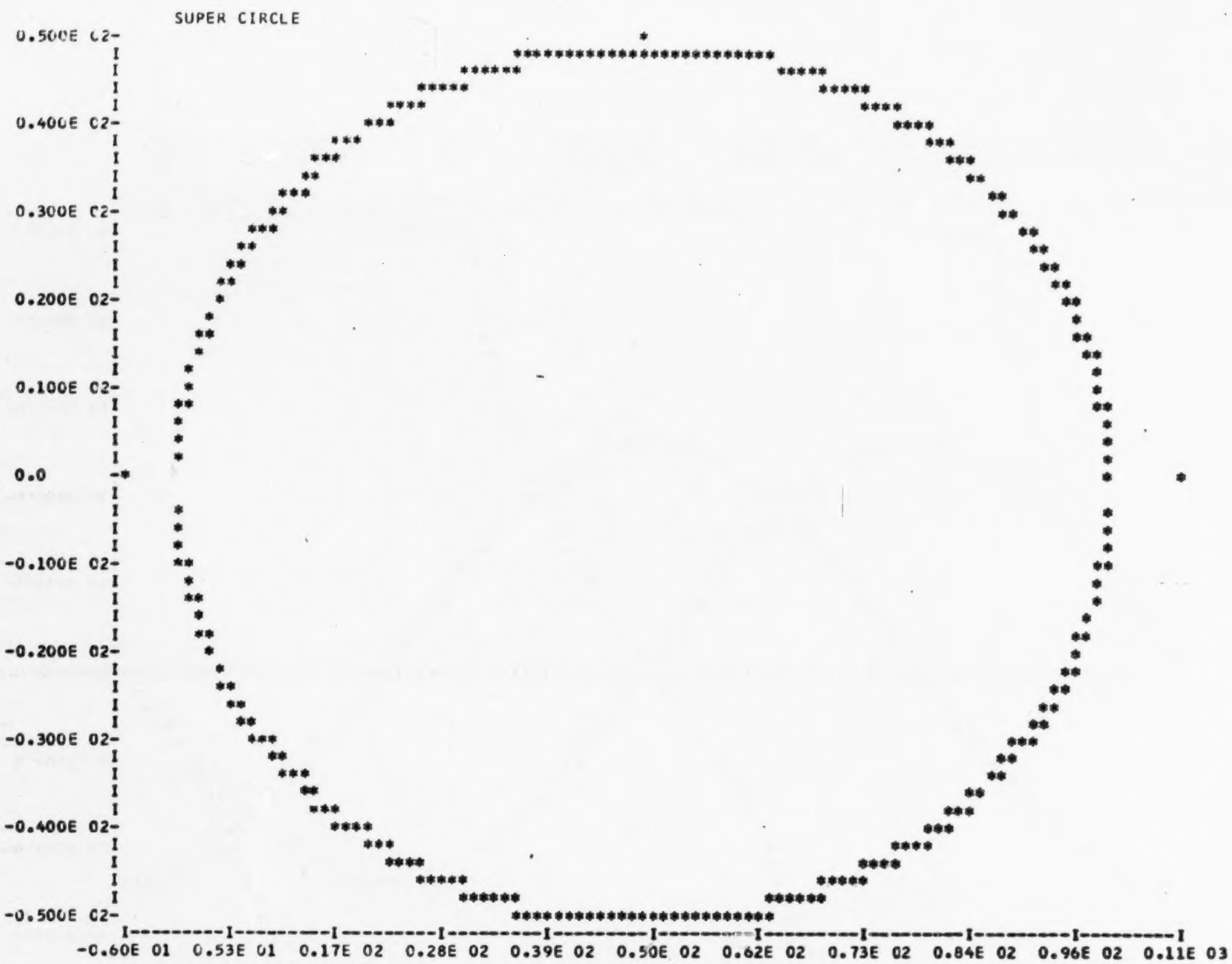
000 0112



```
MAX. X =  0.36000E 03     MAX. Y =  0.20000E 01     SCALE ON X-AXIS =  0.35900E 02 UNITS/INCH
MIN. X =  0.10000E 01     MIN. Y = -0.20000E 01     SCALE ON Y-AXIS =  0.48000E 00 UNITS/INCH
```

continued

000 0112



```
MAX. X =  0.10700E 03       MAX. Y =  0.50000E 02       SCALE ON X-AXIS =  0.11300E 02 UNITS/INCH
MIN. X = -0.60000E 01       MIN. Y = -0.50000E 02       SCALE ON Y-AXIS =  0.12000E 02 UNITS/INCH


TEST OF ERROR ROUTINES

***-ERROR-***    SUBROUTINE LPLOT CALLED WITH IMPROPER PARAMETER LIST
                  NO PLOT GENERATED

***-ERROR-***    SUBROUTINE LPLOT CALLED WITH IMPROPER PARAMETER LIST
                  NO PLOT GENERATED
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $2.37.

Charge to user = computer costs + postage + network overhead
= $2.37 + postage + network overhead

CONTENTS—LPLOT

pages
1 Identification & Abstract
3- 4 User Instructions
5-10 I/O
11 Cost—Contents

DESCRIPTIVE TITLE	Smallest Space Analysis	
CALLING NAME	SSA-1	
INSTALLATION NAME	Office of Data Analysis Research Educational Testing Service	
AUTHOR(S) AND AFFILIATION(S)	Procedure due to:	L. Guttman J. Lingoes University of Michigan
	Program due to:	J. Lingoes University of Michigan
	Adaptation due to:	D. Kirk Educational Testing Service
LANGUAGE	MAP	
COMPUTER	IBM 360/65	
PROGRAM AVAILABILITY	Decks and listings presently available	
CONTACT	Mr. Ernest Anastasio, Off. of Data Anal. Research, Educational Testing Service Rosedale Road, Princeton, N.J. 08540 Tel.: (609) 921-9000 ext. 2552	

FUNCTIONAL ABSTRACT

SSA-1 represents a nonmetric technique for finding the smallest Euclidean space for a configuration of points. To quote from the author's description,

> Briefly stated, the problem posed for the...program is: given a matrix of inequalities among pairs of points in a metric or nonmetric space, determine a set of Euclidean coordinates such that the distances calculated from then are a monotonic function of the ranks or order among the inequalities.

(The voluminous list of references related to this topic are available through the contact person.)

According to the authors, a major computational advantage of this program is that it avoids the local minima problem inherent in the Kruskal program (see EIN No. 000 0068) by employing a rank-image principle. The program can handle up to 70 variables in ten dimensions.

USER INSTRUCTIONS

Input Deck

The input consists merely of three cards plus a Data Deck for each case to be run. The program is written to handle multiple cases. The order of the cards in the input deck are,

Title Card
Parameter Card
Format Card
Data Deck

Title Card

Columns	*Contents*
1	1
2–72	Any title

Parameter Card (Column entries should be right-justified.)

Parameter	*Columns*	*Contents*	
NR	1– 4	Number of variables (≤ 70)	
	5– 8	Minimum number of dimensions	
DMAX	9–12	Maximum number of dimensions (≤ 10)	
	13–16	Blank: 1:	dissimilarities present similarities present
	17–20	1: 0:	distance matrix (derived coefficients) is to be printed distance matrix not printed
	21–24	1: 0:	coordinates are printed coordinates not printed
	25–28	1: 0:	city block model Euclidean model
	29–32	0:	tied blocks not formed; if tied blocks are formed punch width with a decimal for coefficients to be considered tied.

Format Card

Standard FORTRAN format for the following Data Deck: Do not punch the word FORMAT.

continued

Data Deck

This deck must be punched to conform to the Format Card above. It must contain the upper half of a square symmetric matrix without the diagonal elements, in row order. Thus, there will be (NR–1) variables on the first card (or cards) and there will be (NR–1) sets.

System Control Cards

JOB Card will be prepared by ETS personnel.

System Card 1

Columns	*Contents*
1–34	//STEPNAME EXEC GITNGO,NAME=SSA–1

System Card 2

Columns	*Contents*
1–19	//GO.SYSIN DD *

System Card 3

Columns	*Contents*
1– 2	//

Job Deck

JOB Card

System Card 1

System Card 2

Input Deck (may have multiple input decks)

/*
System Card 3

Output

The output consists of

(a) original coefficients (distance)
(b) coordinates for various dimensions
(c) vector plots
(d) derived coefficients (distances)

SAMPLE INPUT

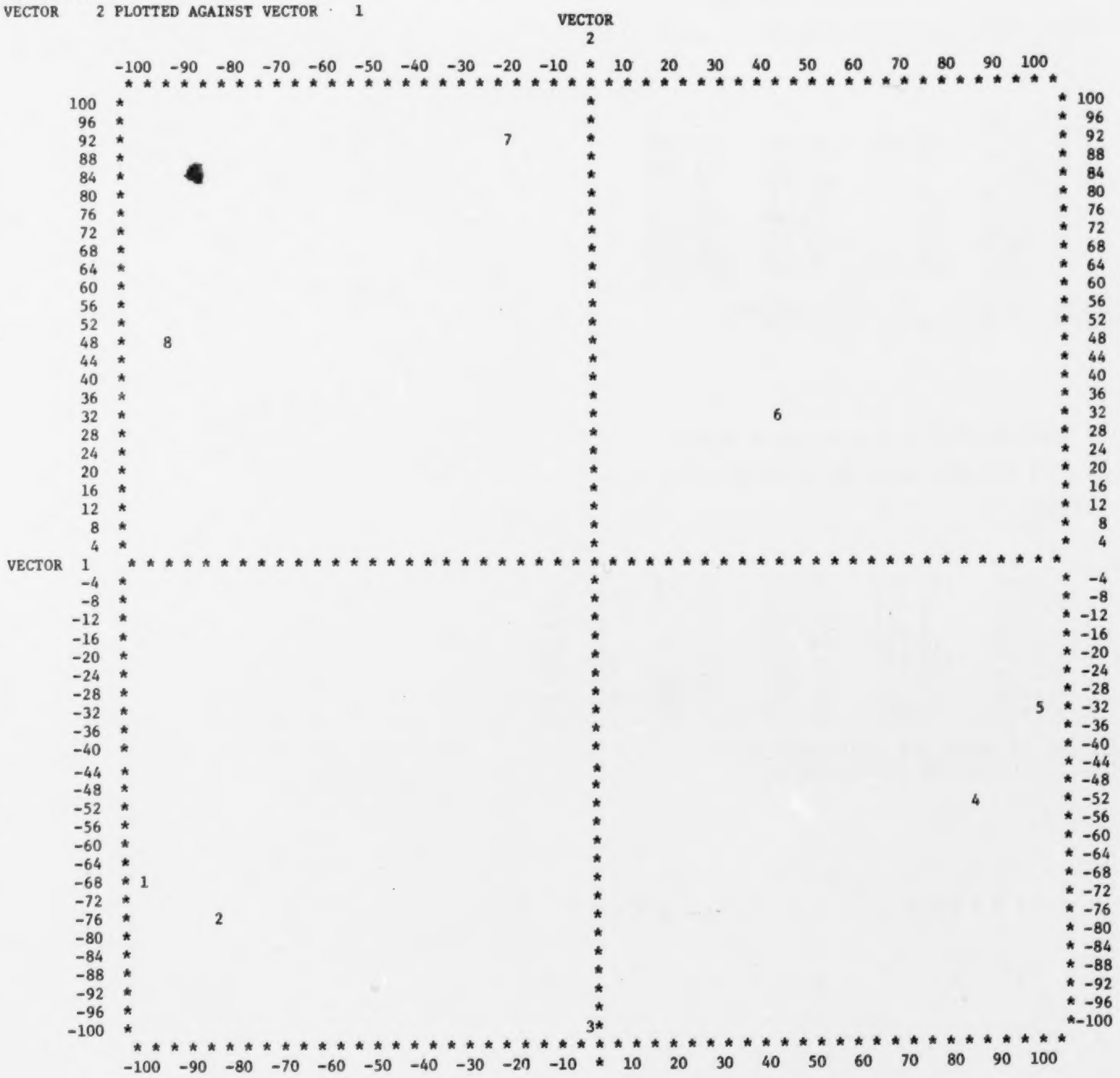
```
//           JOB  (....
//STEPNAME  EXEC   GITNGO,NAME=SSA1
//GO.SYSIN  DD     *
1GUILFORD'S NUMERICAL AND PICTORIAL TESTS (1952)
   8      7   1   1
 (2X35F2.2)
0067401912252639
00502620282638
005239311824
0055492522
00462914
004238
0040
/*
//
```

SAMPLE OUTPUT

```
        O R I G I N A L    C O E F F I C I E N T S

COLUMN =          1     2     3     4     5     6     7     8

 ROW =   1    0.0

 ROW =   2    0.67  0.0

 ROW =   3    0.40  0.50  0.0

 ROW =   4    0.19  0.26  0.52  0.0

 ROW =   5    0.12  0.20  .039  0.55  0.0

 ROW =   6    0.25  0.28  0.31  0.49  0.46  0.0

 ROW =   7    0.26  0.26  0.18  0.25  0.29  0.42  0.0

 ROW =   8    0.39  0.38  0.24  0.22  0.14  0.38  0.40  0.0
```

continued

VECTOR PLOTS
VECTOR 2 PLOTTED AGAINST VECTOR 1



continued

348

```
GUILFORD'S NUMERICAL AND PICTORIAL TESTS (1952).

GUTTMAN-LINGOES SMALLEST SPACE COORDINATES FOR M =  2.

          DIMENSION          1          2
---------------------------------------------------------------

VARIABLE      N. PHI
   1        0.00007  -100.000   -66.561   104.106
   2        0.00005   -80.497   -75.214    91.999
   3        0.00006     3.365  -100.000    80.561
   4        0.00015    80.125   -51.043    92.311
   5        0.00002   100.000   -32.095   107.514
   6        0.00004    43.084    28.822    69.892
   7        0.00009   -13.355    89.411   109.693
   8        0.00010   -87.459    45.998   104.243

NORMALIZED PHI =0.00057 FOR   9 ITERATIONS
COEFFICIENT OF ALIENATION =0.337683E-01



GUILFORD'S NUMERICAL AND PICTORIAL TESTS (1952).

GUTTMAN-LINGOES SMALLEST SPACE COORDINATES FOR M =  3.

          DIMENSION          1          2          3
---------------------------------------------------------------

VARIABLE      N. PHI
   1        0.00000  -100.000   -57.461   -26.269   103.735
   2        0.00000   -75.458   -68.158   -38.187    85.469
   3        0.00000     6.497  -100.000   -54.472    82.016
   4        0.00000    76.695   -49.439   -71.641    91.210
   5        0.00000   100.000   -23.374   -27.717   108.220
   6        0.00000    39.237    31.296   -63.496    69.555
   7        0.00000   -17.232    74.099    -5.199   103.395
   8        0.00000   -78.928    39.495  -100.000   106.803

NORMALIZED PHI =0.00000 FOR  20 ITERATIONS
COEFFICIENT OF ALIENATION =0.598020E-03



             D E R I V E D     C O E F F I C I E N T S


COLUMN =         1     2     3     4     5     6     7     8

 ROW =   1   0.0

 ROW =   2   0.21  0.0

 ROW =   3   1.09  0.87  0.0

 ROW =   4   1.81  1.62  0.91  0.0

 ROW =   5   2.03  1.86  1.18  0.27  0.0

 ROW =   6   1.72  1.62  1.35  0.88  0.83  0.0

 ROW =   7   1.78  1.78  1.90  1.69  1.66  0.83  0.0

 ROW =   8   1.13  1.21  1.72  1.94  2.03  1.32  0.86  0.0
```

COST ESTIMATE

The pricing algorithm for the Educational Testing Service is based on a complex formula involving the amount of core usage, card-processing, tape and dedicated device usage, lines printed and central processor unit (CPU) elapsed time. Computer costs for the job included in the Sample Input were $1.79.

Charge to user = computer costs + postage and handling + network overhead

= $1.79 + $5.00 + network overhead

$6.79 + network overhead

CONTENTS—SSA-1

pages

1 Identification & Abstract
3– 4 User Instructions
5– 8 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Best Least Squares Polynomial Approximation Subroutine
CALLING NAME	LSPOL
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	James Hurt Department of Mathematics University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

The LSPOL routine computes the coefficients of the polynomial of degree n which best approximates (in the least squares sense) a set of data. The degree of the polynomial must be less than or equal to 20. The method used is that of Beckett and Hurt[1].

REFERENCES

1. Beckett, R., and Hurt, J., *Numerical Calculations and Algorithms*, (McGraw-Hill Book Co., Inc., New York, 1967), pp. 264-290.

USER INSTRUCTIONS

Calling Sequence

```
CALL LSPOL1(NMAX)
CALL LSPOL2(W,X,Y,M)
CALL LSPOL3(C,N)
CALL LSPOL4(R)
```

where

NMAX = maximum degree of polynomial desired (≤ 20)

W = name of array containing weights for the data points (there must be M elements in W, each of which is greater than 0)

X = name of array of abscissas

Y = name of array of ordinates

M = number of data points. M is the dimension for the W, X, and Y arrays ($M \geq 1$)

C = the array of calculated coefficients. It will contain N + 1 elements

N = degree of polynomial desired ($1 \leq N \leq NMAX$)

R = calculated index of correlation.

Note: W, X, Y, C, and R *must* be in double precision.

Usage

For each set of data, the following procedure should be followed. First, the sums should be initialized by a call to LSPOL1. The value for NMAX should be no larger than the largest degree polynomial desired. For example, if fourth, fifth, and seventh degree polynomials are desired for this set of data, then NMAX should be 7.

Following the call to LSPOL1, the data points should be read in or computed. Then LSPOL2 can be called to update the sums. Each call to LSPOL2 will update the sums using the M data points given in X and Y. Each data point is weighted in the sums using the values in W. Each value of W should be greater than zero. If all data points are to be weighted equally, a constant can be substituted for W in the call list (i.e. CALL LSPOL2 (1,DO,X,Y,1)). Since consecutive calls to LSPOL2 merely add to each sum without re-initializing, data points can be read into

continued

the computer in groups. For example, if 12,000 data points are used, they may be read in 12 groups of 1000 each by reading in 1000 points, calling LSPOL2 with M = 1000, and repeating this 12 times.

After all data points have been passed to LSPOL2, the coefficients of the desired least squares polynomials can be computed by successive calls to LSPOL3. The resulting least squares polynomial is $p(x) = c_1 + c_2x + c_3x^2 + \ldots + c_{n+1}x^n$. The sums computed by LSPOL2 are used but not destroyed by LSPOL3. Thus, several different degree polynomials can be computed by calling LSPOL3 several different times with no intervening calls to LSPOL2. Calls to LSPOL2 are allowed after calling LSPOL3 and have the effect of adding more data for succeeding calls to LSPOL3.

If the index of correlation is desired, it can be calculated by a call to LSPOL4 immediately after a call to LSPOL3.

Method

LSPOL1 initializes the routines by zeroing the sums. LSPOL2 updates the required sums of powers of X and sums of Y times powers of X. LSPOL3 arranges these sums into an array which represents the equations for the coefficients and then solves these equations. LSPOL4 computes the index of correlation between the data and the most recently computed solution.

Additional Notes

Extreme caution should be used when high degree polynomials are required. The least squares equations for the coefficients become very ill-conditioned, and small round-off errors in the computations may lead to large errors in the computed results. Even if there are no round-off errors, the approximations may not be improved by using higher degree polynomials. See references below for a discussion of this problem.

All computations are done in double precision to reduce the effect of round-off error. The program requires about 6K (K=1000) bytes of core storage plus space for subroutines CROUT, VDOT, and the arrays.

Since LSPOL is a subroutine, the user must embed it in a larger program to provide input and output routines.

REFERENCES

Beckett, R., and Hurt, J., *Numerical Calculations and Algorithms*, (McGraw Hill Book Co., Inc., New York, 1967), pp. 283–286.

Hurt, J., *Some Numerical Techniques for Use on a Digital Computer*, M.S. Thesis, (University of Iowa, 1963), pp. 28–65.

SAMPLE INPUT

```
C
C  EXAMPLE 1
C     THE PROBLEM IS TO APPROXIMATE Y=SQRT(X) USING THE DATA POINTS
C     X=0.1,0.2,...,1.0)
      REAL*8 X(1),Y(1),C(5),R,X1,Y1,Z,Y2,E,E2
      WRITE(6,1000)
1000  FORMAT('1','EXAMPLE 1 - A=SQRT(B)')
C     INITIALIZE BY CALLING LSPOL1
      CALL LSPOL1(4)
      DO 1 J=1,10
C     COMPUTE POINT FOR X=X(J)
      X(1)=0.1D0*DBLE(FLOAT(J))
      Y(1)=DSQRT(X(1))
C     UPDATE SUMS BY CALLING LSPOL2 (ALL WEIGHTS =1.0)
    1 CALL LSPOL2(1.D0,X,Y,1)
C     CONTROL TO SOLVE FOR SEVERAL DIFFERENT DEGREE POLYNOMIALS
      DO 2 N=1,4
C     SOLVE FOR N-TH DEGREE POLYNOMIAL
      CALL LSPOL3(C,N)
C     COMPUTE CORRELATION COEFFICIENT
      CALL LSPOL4(R)
C     WRITE SOLUTION
    2 WRITE(6,3) N,R,C(1),(C(I+1),I=1,N)
    3 FORMAT(1H0,I5,6D20,8)
      X1=0.1D0
      DO 4 I=1,10
      Y1=DSQRT(X1*DBLE(FLOAT(I)))
      Z=X1*DBLE(FLOAT(I))
      Y2=C(1)+C(2)*Z+C(3)*Z**2+C(4)*Z**3+C(5)*Z**4
      E=DABS(Y2-Y1)
      E2=E/Y1
    4 WRITE(6,6) Y1,Y2,E,E2
    6 FORMAT(' ',4D20.8)
      CALL EXIT
      END
/*
```

```
C
C  EXAMPLE 2
C  LEAST SQUARES POLYNOMIAL
      REAL*8 W(6),X(6),Y(6),C(6),R
      WRITE(6,1000)
1000  FORMAT('1','EXAMPLE 2 - LEAST SQUARES POLYNOMIAL')
      WRITE(6,2000)
2000  FORMAT('0',2X,'DATA POINTS'/'0',4X,'X',7X,'Y')
      CALL LSPOL1(5)
      DO 1 N=1,6
      W(N)=1.D0
      READ(5,3000)X(N),Y(N)
3000  FORMAT(D5.0,D10.0)
      I=X(N)
      J=Y(N)
    1 WRITE(6,4000) I,J
4000  FORMAT(' ',I5,I10)
      WRITE(6,5000)
5000  FORMAT('0',1X,'N',4X,'CORRELATION',51X,'COEFFICIENTS')
      CALL LSPOL2(W,X,Y,6)
      DO 2 N=1,5
      CALL LSPOL3(C,N)
      CALL LSPOL4(R)
      M=N+1
    2 WRITE(6,6000)N,R,(C(J),J=1,M)
6000  FORMAT(1H0,I2,8D16.8)
      CALL EXIT
      END
/*
//GO.SYSIN DD *
    1        25
    2       168
    3      1051
    4       348
    5     13869
    6     35200
/*
```

SAMPLE OUTPUT

```
EXAMPLE 1 - A =SQRT(B)
    1      0.98919384D 00     0.30791386D 00     0.73199179D 00
    2      0.99915172D 00     0.21731804D 00     0.11849709D 01     -0.41179918D 00
    3      0.99990938D 00     0.17693507D 00     0.15431461D 01     -0.11883946D 01     0.47066391D 00
    4      0.99998991D 00     0.15412663D 00     0.18355620D 01     -0.22583710D 01     0.19327434D 01     -0.66458160D 00
   0.31622777D 00     0.31696541D 00     0.73764432D-03     0.23326362D-02
   0.44721360D 00     0.44530282D 00     0.19107805D-02     0.42726350D-02
   0.54772256D 00     0.54834281D 00     0.62025508D-03     0.11324257D-02
   0.63245553D 00     0.63369437D 00     0.12388420D-02     0.19587811D-02
   0.70710678D 00     0.70737147D 00     0.26469323D-03     0.37433276D-03
   0.77459667D 00     0.77379309D 00     0.80357603D-03     0.10374122D-02
   0.83666003D 00     0.83578321D 00     0.87681267D-03     0.10479916D-02
   0.89442719D 00     0.89457082D 00     0.14363301D-03     0.16058658D-03
   0.94868330D 00     0.94978992D 00     0.11066174D-02     0.11664772D-02
   0.10000000D 01     0.99947948D 00     0.52051584D-03     0.52051584D-03
```

```
EXAMPLE 2 - LEAST SQUARES POLYNOMIAL
  DATA POINTS
    X       Y
    1        25
    2       168
    3      1051
    4       348
    5     13869
    6     35200
 N   CORRELATION                                              COEFFICIENTS
 1  0.81545231D 00 -0.13184000D 05  0.61792857D 04
 2  0.97728657D 00  0.12898000D 05 -0.13382214D 05  0.27945000D 04
 3  0.99648055D 00 -0.64173333D 04  0.10915249D 05 -0.52535556D 04  0.76648148D 03
 4  0.99677853D 00 -0.12549333D 05  0.21475915D 05 -0.10935389D 05  0.19588148D 04 -0.85166667D 02
 5  0.10000000D 01  0.61444000D 05 -0.13517000D 06  0.10481933D 06 -0.36532666D 05  0.58006666D 04 -0.33633333D 03
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for each of the two examples in the Sample Output above was approximately $1.30.

Charge to user = computer costs + postage + network overhead
= $2.60 + postage + network overhead

CONTENTS—LSPOL

pages	
1	Identification & Abstract
3– 4	User Instructions
5– 6	I/O
7	Cost—Contents

DESCRIPTIVE TITLE	Stepwise Regression Analysis
CALLING NAME	STEPREGN
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Louise R. Levine The University of Iowa Computer Center Revision of BMD02R program
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

STEPREGN computes a sequence of multiple linear regression equations in a stepwise manner. At each step one variable is added to the regression equation. The variable added is the one which makes the greatest reduction in the error sum of squares. In addition, it is the variable which has highest partial correlation with the dependent variable partialed on the variables which have already been added; and it is the variable which, if it were added, would have the highest F value. Variables can also be forced into the regression equation. Non-forced variables are automatically removed when their F values become too low. Regression equations with or without the regression intercept may be selected.

Limitations per problem:

p	number of original variables ($2 \le p \le 80$)
q	number of variables added by transgeneration ($-9 \le q \le 78$)
$p+q$	total number of variables ($2 \le p+q \le 80$)
s	number of Sub-problem Cards ($1 \le s \le 99$)
k	number of Variable Format Cards ($1 \le k \le 10$)
i	number of variables to be plotted ($0 \le i \le 30$)
n	number of cases ($1 \le n \le 9999$)
m	number of Transgeneration Cards ($0 \le m \le 99$)

continued

Estimation of output pages per problem.

Number of pages = $\frac{\text{no. of steps } [23 + (3/4)(p+q)]}{56}$ + 5 per sub-problem.

Transgeneration of the variables is allowed (see Transgeneration Codes).

REFERENCE

Dixon, W.J., (Ed.), *BMD: Biomedical Computer Programs*, (Univ. of Calif. Press, Berkeley, Calif., 1967), pp. 233-259d.

USER INSTRUCTIONS

Input Deck

Problem Card (One Problem Card for each problem)

Columns	*Contents*
1– 6	PROBLM
10–15	Alphanumeric problem name
17–20	Sample size ($1 \leq n \leq 9999$)
24–25	Number of original variables ($2 \leq p \leq 80$)
29–30	Number of Transgeneration Cards ($0 \leq m \leq 99$)
34–35	Number of variables added by transgeneration ($-9 \leq q \leq 78$)
39–40	Tape number if data are on tape (≠ logical 1 or 2) Blank if data are not on tape.
44–45	Number of Sub-problem Cards ($1 \leq s \leq 99$)
48–49	Number of variables labeled on Labels Cards. Leave blank if Labels Cards are not used.
51–53	YES: If means and standard deviations are to be printed; otherwise, leave blank.
55–57	YES: If covariance matrix is to be printed; otherwise, leave blank.
59–61	YES: If correlation matrix is to be printed; otherwise, leave blank.
63–65	YES: If zero regression intercept is desired; otherwise, leave blank.
68–69	NO: If tape specified in Col. 39–40 is not to be rewound before this problem; leave blank if Col. 39–40 are blank, or if tape rewind is desired.
71–72	Number of F-type Variable Format Cards ($1 \leq k \leq 10$)

Transgeneration Cards (Optional)

Columns	*Contents*
1– 6	TRNGEN
7– 9	Variable index k
10–11	Code from transgeneration list

continued

Transgeneration Cards *(continued)*

Columns	*Contents*
12–14	Variable index i
15–20	Variable index j or constant c (The constant c is punched with a decimal point if used with variables which have an F-type format.)

Transgeneration List

Notation to be used in the following transgeneration list:

i,j,k are variable indices (need not be different),
c is a constant,
n is the number of cases, or sample size.

Code	*Transgeneration*	*Restriction*
01	$\sqrt{X_i} \longrightarrow X_k$	$X_i \geq 0$
02	$\sqrt{X_i} + \sqrt{X_i + 1} \longrightarrow X_k$	$X_i \geq 0$
03	$\log_{10} X_i \longrightarrow X_k$	$X_i > 0$
04	$e^{X_i} \longrightarrow X_k$	
05	$\arcsin \sqrt{X_i} \longrightarrow X_k$	$0 \leq X_i \leq 1$
06	$\arcsin \sqrt{X_i/(n+1)} + \arcsin \sqrt{(X_i+1)/(n+1)} \longrightarrow X_k$	$0 \leq (X_i/n) \leq 1$
07	$1/X_i \longrightarrow X_k$	$X_i \neq 0$
08	$X_i + c \longrightarrow X_k$	
09	$X_i X c \longrightarrow X_k$	
10	$X_i^c \longrightarrow X_k$	$X_i \geq 0$
11	$X_i + X_j \longrightarrow X_k$	
12	$X_i - X_j \longrightarrow X_k$	
13	$X_i X_j \longrightarrow X_k$	
14	$X_i / X_j \longrightarrow X_k$	

continued

Code	*Transgeneration*	*Restriction*
15	If $X_i \geq c$, $1 \rightarrow X_k$; otherwise, $0 \rightarrow X_k$	
16	If $X_i \geq X_j$, $1 \rightarrow X_k$; otherwise, $0 \rightarrow X_k$	
17	$\log_e X_i \rightarrow X_k$	$X_i > 0$
20	$\sin X_i \rightarrow X_k$	
21	$\cos X_i \rightarrow X_k$	
22	$\arctan X_i \rightarrow X_k$	
23	$X_i^{X_j} \rightarrow X_k$	$X_i > 0$
24	$c^{X_i} \rightarrow X_k$	$c > 0$

Labels Cards (Optional)

Labels Cards allow the user to substitute alphanumeric names for the usual numeric indices (variable numbers or category designations) which appear on the printed output.

Columns	*Contents*
1– 6	LABELS
7–10	The number of the variable (or category, or index) to be named. This number must be right-justified.
11–16	The corresponding alphanumeric name.
17–20	The number of another variable.
21–26	The corresponding alphanumeric name.
⋮	⋮
67–70	The number of another variable.
71–76	The corresponding alphanumeric name of that variable (up to 7 per card).

There may be from one to seven pairs of variable numbers and labels on each Labels Card. If desired, only one pair may be specified on each card. However, the total number of labels appearing on all the Labels Cards must equal the number of labels specified on the Problem or Sub-problem Card.

continued

It is not necessary to label all the variables. Those labeled may be listed in any order.

Example: Suppose the number of variables to be labeled as specified on the Problem Card is 9. Then the Labels Cards might be punched as:

```
LABELS  10HEIGHT  07WEIGHT  105AGE    003 X1    0051VAR59 0073 X+Y
LABELS  99SEX   0100ANYNAM
LABELS  05STATUS
```

F-type Variable Format Cards

Standard FORTRAN format statements are used to read the Data Cards one case at a time across all variables.

Data Input Deck (If data are on cards)

Arrange the data by case for all variables.

Sub-problem Card

Columns	*Contents*	
1– 6	SUBPRO	
9–10	Number of the dependent variable.	
13–15	Maximum number of steps. This will be 2(p+q) if left blank.	
20–25	F-level for inclusion. This will be 0.01 if left blank.	These are the tabled F values for 1 to ∞ d.f.
30–35	F-level for deletion. This will be 0.005 if left blank.	
39–45	Tolerance level (lowest level you will allow for the partial correlation coefficient). This will be 0.001 if left blank.	
49–50	Number of variables on the Index-Plot Card $(0 \leq i \leq 30)$.	
53–55	YES: Control-Delete Cards are included.	
58–60	YES: List of summary residuals is to be printed.	
63–65	YES: Summary table is to be printed.	

Control-Delete Card (Optional)

Columns	*Contents*
1– 6	CONDEL

continued

Control-Delete Card *(continued)*

Columns	*Contents*
7	Control Value for first variable.
8	Control Value for second variable.
⋮	⋮
72	Control Value for 66th variable.

If there are more than 66 variables, continue on another card of the same form, until p+q variables have been specified.

The variable numbers above refer to variables after transgeneration.

Control Values

Code	*Value*
1	Delete variable (or dependent variable)
2	Free variable
3	Low-level forced variable
⋮	⋮
9	High-level forced variable

If no Control-Delete Cards are included, or if a field is left blank on the Control-Delete Cards included in the deck, the value 2 will be assigned if the variable is not the dependent variable and the value 1 assigned if it is the dependent variable.

Index-Plot Card (Optional)

Variables specified on this card are plotted against the residuals.

Columns	*Contents*
1– 6	IDXPLT
7– 8	First variable to be plotted.
9–10	Second variable to be plotted.
⋮	⋮
65–66	30th variable to be plotted.

continued

No more than 30 variables may be plotted per sub-problem.

Variables specified refer to the original data after transgeneration.

Finish Card

Columns	*Contents*
1– 6	FINISH

Multiple Problems

The Sub-problem Card through the Index-Plot Card may be repeated as many as 99 times each problem. The Problem Card through the Index-Plot Card may be repeated as often as desired. The FINISH Card follows the *last* problem.

Sample Deck Set Up

Problem Card
Transgeneration Card (Optional)
Labels Card (Optional)
F-type Variable Format Card
Data Input Cards (If data are on cards)
Sub-problem Card
Control-Delete Card (Optional)
Index-Plot Card (Optional)
Finish Card
/*

SAMPLE INPUT

```
PROBLM    TSTC2R    68    6    0    0         1   4 YES YES YES  NO      1
LABELS    1  AGE      4HEIGHT     2WEIGHT     5STATUS
(F6.2,F6.0,2F6.2,2F6.0)
 00250 00025 02500 00150 00034 00064
 01300 00021 02100 00087 00036 00065
 00350 00022 02200 00043 00041 00082
 00175 00009 00130 00180 00015 00023
 00300 00023 02300 00200 00033 00064
 00200 00010 00060 00330 00013 00016
 00550 00007 00140 00340 00016 00012
 00600 00006 00080 00500 00011 00027
 00130 00008 00270 00150 00019 00048
 00500 00018 00360 00180 00027 00050
 00500 00003 00100 00140 00014 00012
 00300 00008 00270 00100 00025 00013
 00200 00006 00300 00155 00021 00020
 00200 00008 00100 00250 00018 00023
 00100 00022 02200 00110 00046 00118
 00400 00013 01300 00280 00017 00050
 00050 00026 00120 00073 00048 00063
 00025 00023 02300 00010 00036 00150
 01400 00003 00100 00350 00005 00072
 00250 00015 00250 00028 00033 00054
 00350 00028 01400 00001 00046 00109
 00350 00006 00060 00500 00010 00010
 00250 00035 03500 00570 00038 00125
 00350 00011 00200 00340 00016 00044
 00200 00011 01100 00050 00020 00048
 00700 00032 03200 00660 00038 00105
 00400 00008 00100 00450 00012 00009
 01500 00023 02300 00015 00049 00130
 00100 00038 03800 00220 00043 00160
 00350 00015 00500 00150 00033 00046
 01300 00006 00120 00370 00009 00036
 00200 00025 02500 00100 00035 00150
 01200 00005 00170 00030 00021 00078
 00400 00009 00075 00190 00017 00023
 00300 00007 00350 00260 00012 00042
 00800 00020 02000 00220 00030 00072
 00900 00006 00086 00250 00015 00020
 00600 00012 00400 00120 00020 00036
 00800 00026 00160 00110 00035 00056
 00150 00015 00300 00160 00029 00036
 00700 00010 00090 01000 00012 00026
 00800 00028 02800 00420 00040 00108
 00200 00034 03400 00090 00042 00106
 00600 00004 00080 00360 00011 00016
 01500 00032 03200 00180 00044 00104
 01700 00011 01100 00230 00014 00047
 01600 00002 00050 00180 00011 00027
 00300 00018 00160 00110 00032 00012
 00600 00003 00040 00130 00015 00007
 01400 00008 00110 00200 00017 00018
 00600 00014 00090 00070 00029 00028
 00180 00012 00240 00150 00021 00025
 01500 00003 00150 00080 00013 00011
 01800 00006 00550 00570 00009 00020
 00500 00012 00200 00410 00016 00014
 03000 00011 01100 00200 00022 00038
 02900 00008 00800 00100 00022 00103
 00180 00024 02400 00110 00038 00106
 01300 00026 02600 00170 00038 00063
 01900 00029 02900 04800 00029 00208
 01100 00017 01700 00160 00025 00032
 01000 00015 00500 00350 00019 00028
 00600 00010 00500 00100 00026 00032
 00500 00022 02200 00120 00039 00100
 00100 00015 00500 00080 00029 00050
 01700 00009 00300 01300 00010 00080
 00500 00030 03500 00090 00058 00065
 00130 00010 00130 00900 00010 00025
SUBPRO   6   6     0.500      0.300                3  YES  YES  YES
CONDEL1
IDXPLT 2 3
FINISH
/*
//
```

SAMPLE OUTPUT

```
BMD02R - STEPWISE REGRESSION - REVISED JULY 18, 1968
HEALTH SCIENCES COMPUTING FACILITY, UCLA

PROBLEM CODE                   TST02R
NUMBER OF CASES                    68
NUMBER OF ORIGINAL VARIABLES        6
NUMBER OF VARIABLES ADDED           0
TOTAL NUMBER OF VARIABLES           6
NUMBER OF SUB-PROBLEMS              1
THE VARIABLE FORMAT IS   (F6.2,F6.0,2F6.2,2F6.0)




  VARIABLE         MEAN       STANDARD DEVIATION
    AGE   1       6.99558            6.47375
  WEIGHT  2      15.25000            9.35750
          3      10.42512           11.62698
  HEIGHT  4       3.09955            5.99713
  STATUS  5      25.39705           12.47935
          6      56.79411           43.54991




COVARIANCE MATRIX

VARIABLE      1          2          3          4          5          6
 NUMBER

    1      41.909    -10.690      0.387      9.919    -15.809     23.134
    2                 87.563     94.437      5.652    102.660    305.484
    3                           135.187      8.776    109.104    398.026
    4                                       35.966    -10.510     89.470
    5                                                 155.734    350.634
    6                                                           1896.595




CORRELATION MATRIX

VARIABLE      1          2          3          4          5          6
 NUMBER

    1       1.000     -0.176      0.005      0.255     -0.196      0.082
    2                  1.000      0.868      0.101      0.879      0.750
    3                             1.000      0.126      0.752      0.786
    4                                        1.000     -0.140      0.343
    5                                                   1.000      0.645
    6                                                              1.000




SUB-PROBLM    1
 DEPENDENT VARIABLE              6
 MAXIMUM NUMBER OF STEPS         6
 F-LEVEL FOR INCLUSION     0.500000
 F-LEVEL FOR DELETION      0.300000
 TOLERANCE LEVEL           0.001000
```

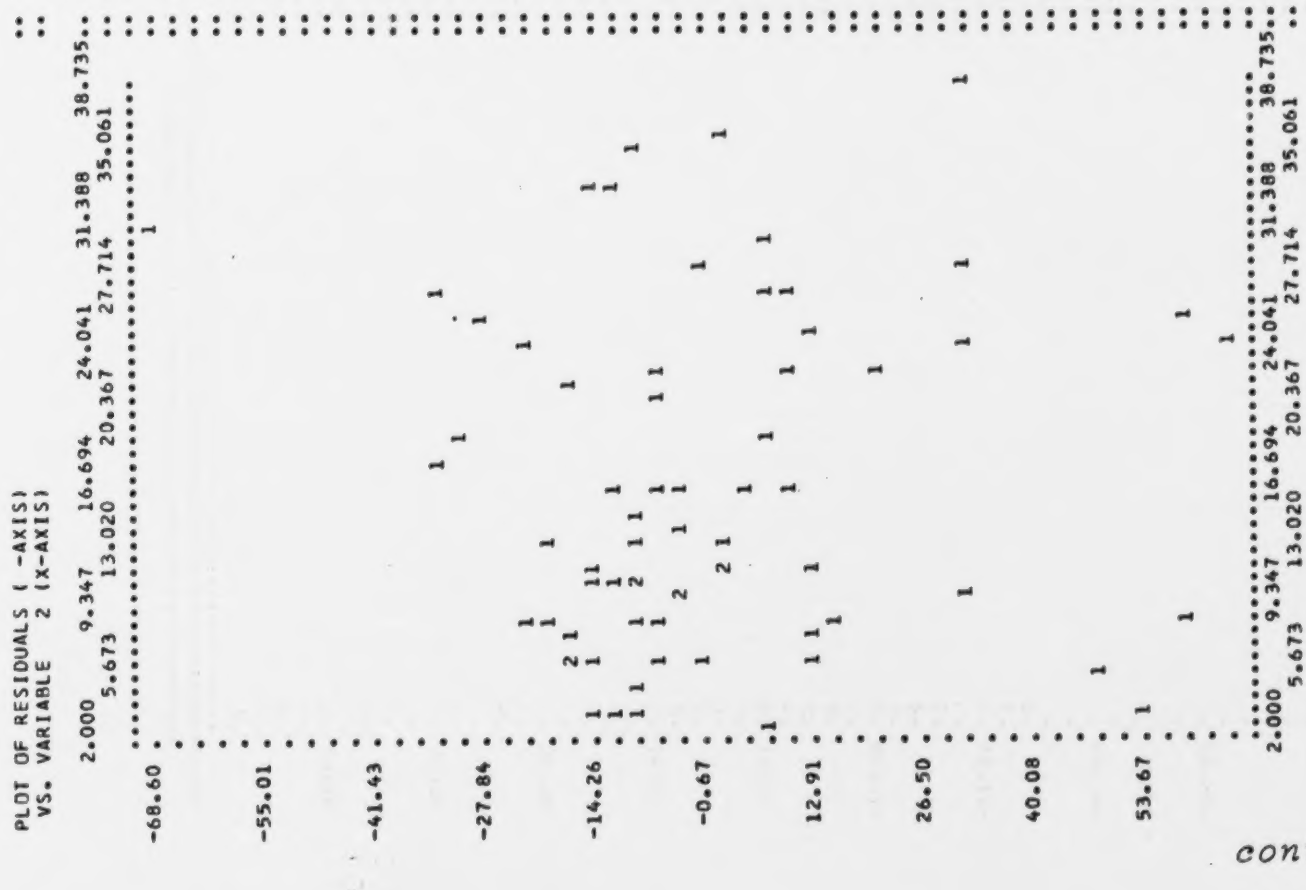
continued

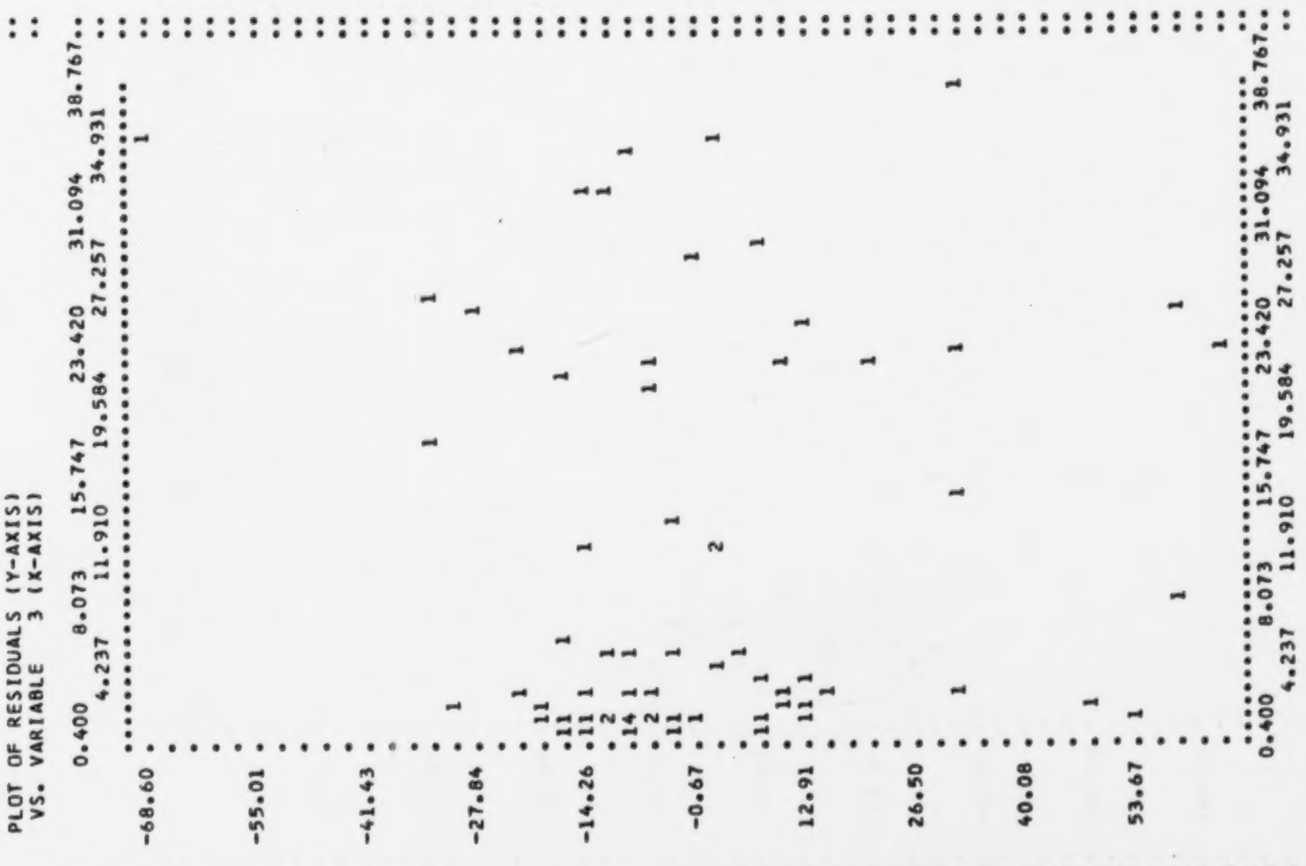
```
 STEP NUMBER    1
VARIABLE ENTERED    3

MULTIPLE R                0.7861
STD. ERROR OF EST.       27.1234

ANALYSIS OF VARIANCE
                       DF    SUM OF SQUARES    MEAN SQUARE    F RATIO
       REGRESSION       1       78517.062       78517.062    106.728
       RESIDUAL        66       48554.734         735.678

                                                           .
                 VARIABLES IN EQUATION                     .                   VARIABLES NOT IN EQUATION
                                                           .
  VARIABLE      COEFFICIENT  STD. ERROR  F TO REMOVE       .       VARIABLE    PARTIAL CORR.     TOLERANCE    F TO ENTER
                                                           .
                                                           .
  (CONSTANT         26.09973 )                             .
       3             2.94427      0.28500    106.7276 (2)  .        AGE    1        0.12622         1.0000        1.0523 (1)
                                                           .     WEIGHT    2        0.21933         0.2466        3.2847 (2)
                                                           .     HEIGHT    4        0.39729         0.9842       12.1823 (2)
                                                           .     STATUS    5        0.13276         0.4346        1.1662 (2)



 STEP NUMBER    2
VARIABLE ENTERED    4

MULTIPLE R                0.8235
STD. ERROR OF EST.       25.0817

ANALYSIS OF VARIANCE
                       DF    SUM OF SQUARES    MEAN SQUARE    F RATIO
       REGRESSION       2       86180.875       43090.437     68.496
       RESIDUAL        65       40890.934         629.091

                                                           .
                 VARIABLES IN EQUATION                     .                   VARIABLES NOT IN EQUATION
                                                           .
  VARIABLE      COEFFICIENT  STD. ERROR  F TO REMOVE       .       VARIABLE    PARTIAL CORR.     TOLERANCE    F TO ENTER
                                                           .
                                                           .
  (CONSTANT         21.74443 )                             .
       3             2.82756      0.26566    113.2883 (2)  .        AGE    1        0.02724         0.9340        0.0475 (1)
 HEIGHT  4           1.79767      0.51504     12.1823 (2)  .     WEIGHT    2        0.24653         0.2465        4.1414 (2)
                                                           .     STATUS    5        0.32179         0.3784        7.3926 (2)
```

continued

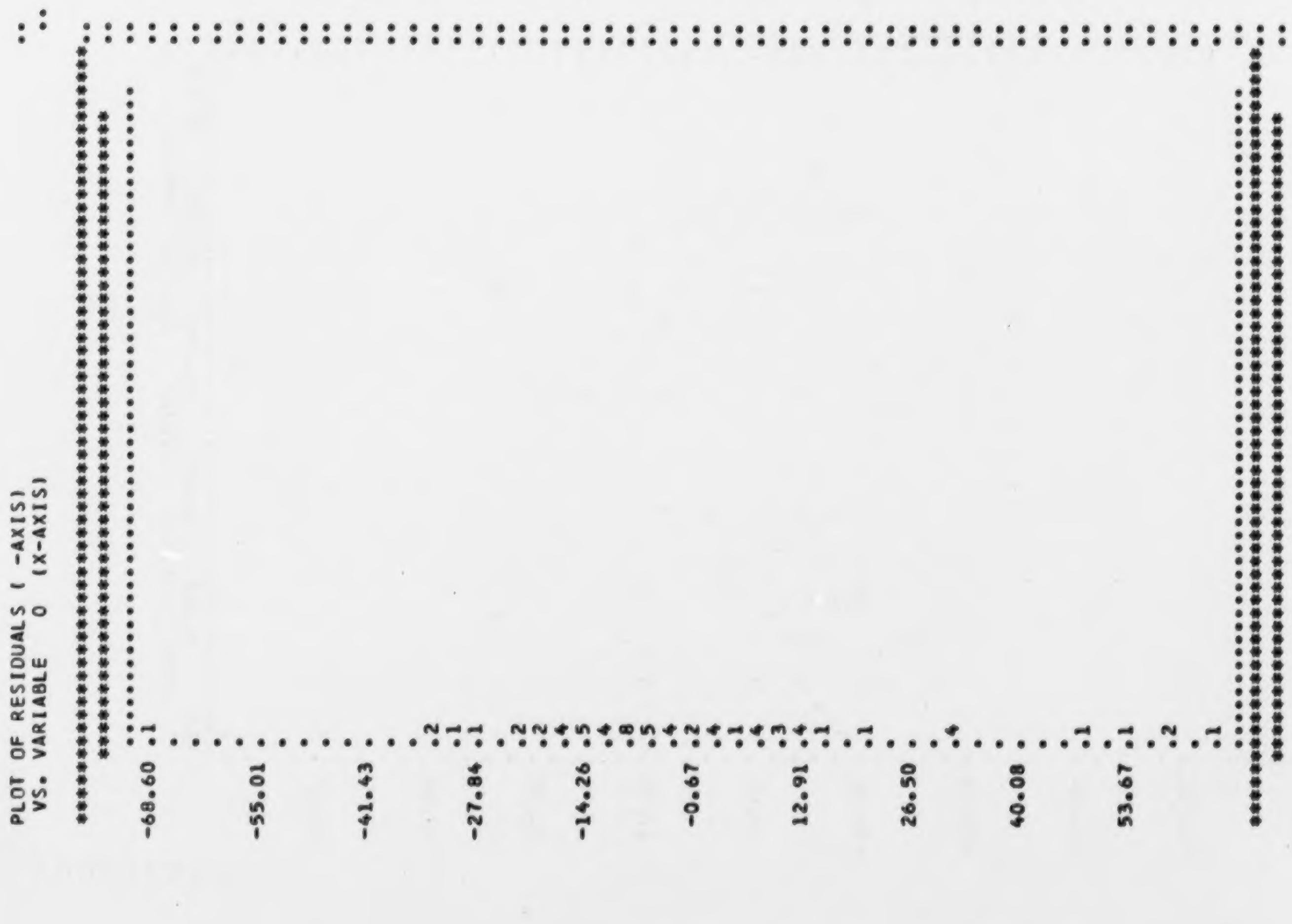
```
 STEP NUMBER    3
VARIABLE ENTERED    5

MULTIPLE R              0.8435
STD. ERROR OF EST.     23.9324

ANALYSIS OF VARIANCE
                       DF    SUM OF SQUARES    MEAN SQUARE    F RATIO
        REGRESSION      3       90415.062       30138.352     52.619
        RESIDUAL       64       36656.719         572.761

                                                          .
             VARIABLES IN EQUATION                        .              VARIABLES NOT IN EQUATION
                                                          .
  VARIABLE      COEFFICIENT  STD. ERROR  F TO REMOVE      .   VARIABLE    PARTIAL CORR.    TOLERANCE    F TO ENTER
                                                          .
                                                          .
 (CONSTANT          2.91084 )                             .
         3          1.95842     0.40797     23.0441 (2)   .    AGE    1       0.11112        0.8832       0.7876 (1)
 HEIGHT  4          2.31237     0.52664     19.2789 (2)   .   WEIGHT  2       0.01574        0.1134       0.0156 (2)
 STATUS  5          1.03552     0.38085      7.3926 (2)   .

F-LEVEL OR TOLERANCE INSUFFICIENT FOR FURTHER COMPUTATION

SUMMARY TABLE
  STEP            VARIABLE                 MULTIPLE                       INCREASE         F VALUE TO      NUMBER OF INDEPENDENT
 NUMBER        ENTERED  REMOVED          R            RSQ                  IN RSQ        ENTER OR REMOVE    VARIABLES INCLUDED

   1                  3                0.7861       0.6179                 0.6179          106.7276                 1
   2           HEIGHT 4                0.8235       0.6782                 0.0603           12.1823                 2
   3           STATUS 5                0.8435       0.7115                 0.0333            7.3926                 3


  LIST OF RESIDUALS

 CASE          Y               Y
 NUMBER      X( 6)         COMPUTED       RESIDUAL

    1       64.0000         90.5476       -26.5476
    2       65.0000         83.3282       -18.3282
    3       82.0000         89.4467        -7.4467
    4       23.0000         25.1519        -2.1519
    5       64.0000         86.7514       -22.7514
    6       16.0000         25.1785        -9.1785
    7       12.0000         30.0830       -18.0830
    8       27.0000         27.4302        -0.4302
    9       48.0000         31.3420        16.6580
   10       50.0000         42.0825         7.9175
    .          .               .              .
    .          .               .              .
    .          .               .              .
   60      208.0000        200.7290         7.2710
   61       32.0000         65.7918       -33.7918
   62       28.0000         40.4711       -12.4711
   63       32.0000         41.9388        -9.9388
   64      100.0000         89.1563        10.8437
   65       50.0000         44.5829         5.4171
   66       80.0000         49.2021        30.7979
   67       65.0000        133.5969       -68.5969
   68       25.0000         36.6233       -11.6233
```

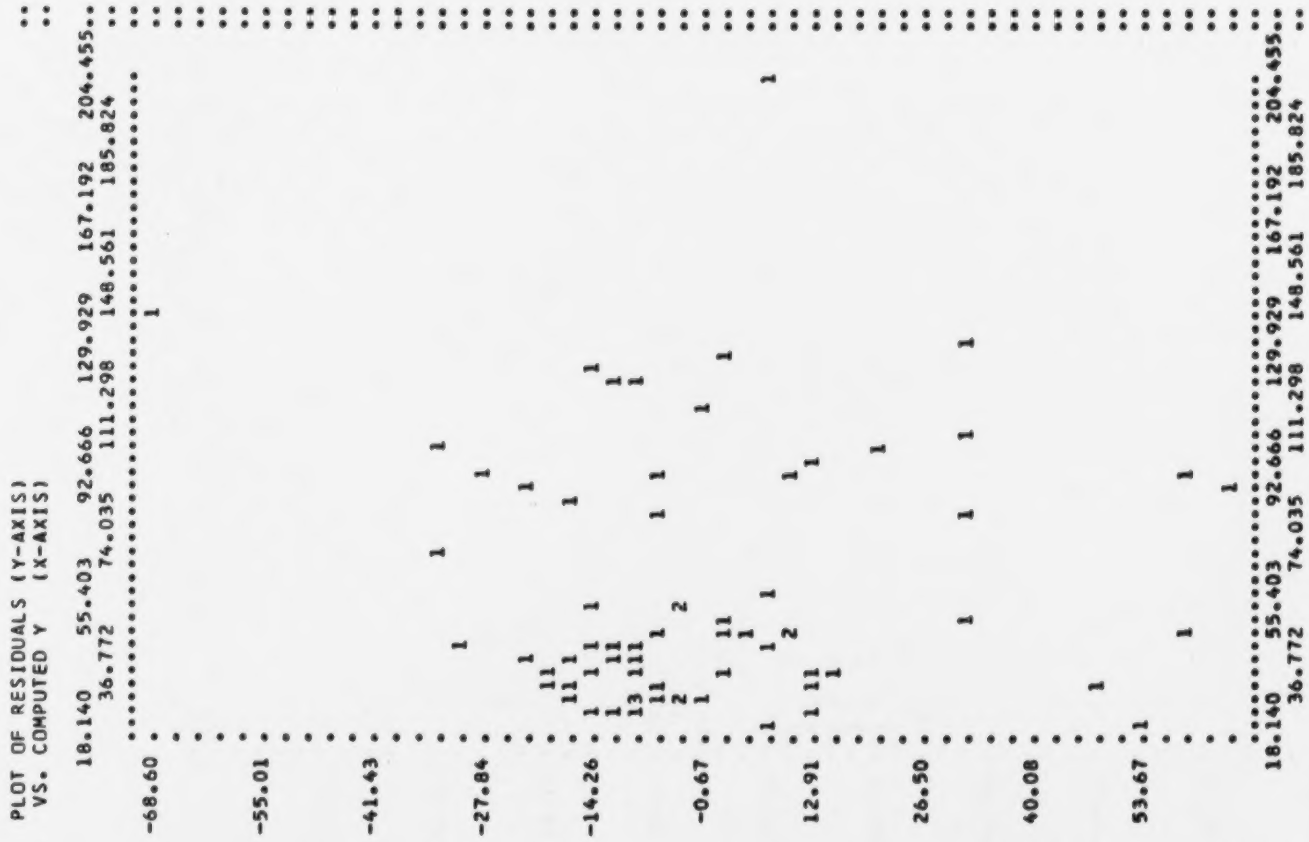
continued


PLOT OF RESIDUALS (-AXIS)
VS. VARIABLE 2 (X-AXIS)
2.000
5.673
9.347
13.020
16.694
20.367
24.041
27.714
31.388
35.061
38.735
-68.60
-55.01
-41.43
-27.84
-14.26
-0.67
12.91
26.50
40.08
53.67


PLOT OF RESIDUALS (Y-AXIS)
VS. VARIABLE 3 (X-AXIS)
0.400
4.237
8.073
11.910
15.747
19.584
23.420
27.257
31.094
34.931
38.767
-68.60
-55.01
-41.43
-27.84
-14.26
-0.67
12.91
26.50
40.08
53.67

continued


PLOT OF RESIDUALS (-AXIS)
VS. VARIABLE 0 (X-AXIS)
-68.60
-55.01
-41.43
-27.84
-14.26
-0.67
12.91
26.50
40.08
53.67


PLOT OF RESIDUALS (Y-AXIS)
VS. COMPUTED Y (X-AXIS)
18.140
55.403
92.666
129.929
167.192
204.455
36.772
74.035
111.298
148.561
185.824
-68.60
-55.01
-41.43
-27.84
-14.26
-0.67
12.91
26.50
40.08
53.67

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $1.14.

Charge to user = computer costs + postage + network overhead
= $1.14 + postage + network overhead

CONTENTS—STEPREGN

pages

1– 2 Identification & Abstract
3– 8 User Instructions
9–14 I/O
15 Cost—Contents

DESCRIPTIVE TITLE	General Program for Multivariate Cross-classification
CALLING NAME	NUCROS
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	K. Janda Northwestern University Adapted for IBM 7044: Merle Wood, Political Science Department, University of Iowa Converted for IBM 360: Janice Lewis, University Computer Center, University of Iowa
LANGUAGE	FORTRAN IV (G)
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

NUCROS prepares multivariate cross-classifications in accordance with user-supplied instructions. It will compute row and column sums, and if the user desires, it will also compute row and column percentages, the chi square value, the contingency coefficient, and the Kendall tau-b or tau-c, the Goodman-Kruskal gamma, and Somers D_{yx} and D_{xy} correlation coefficients. There may be up to 40 variables and up to 72 tables processed for each cycle, with a maximum of 99 cycles. Two variables are cross-classified in each table, but a third and fourth variable may be held constant. The data are assumed to be positive integers.

REFERENCES

Janda, K., *Data Processing*, (Univ. of Northwestern Press, Evanston, Ill., 1964).

USER INSTRUCTIONS

Input Deck

The user must supply as his data the following cards.

Cycle Card
Eleven Instruction Cards
Recoding Control Cards (optional)
Variable Identification Cards (optional)
Raw Data Cards unless tape input is desired

may be repeated for additional cycles

Cycle Card

The number punched in Col. 1–2 of the Cycle Card tells the main program how many times it will be executed. For normal runs only one cycle is needed, and 01 should be punched in Col. 1–2. However, if more than 40 variables must be processed or more than 72 tables constructed, the number of cycles may be increased to a maximum of 99 with the desired number of cycles punched in Col. 1–2.

Eleven Instruction Cards for each cycle

Title Cards (1 and 2)

Columns	*Contents*
Card 1, 1–72	Any title information desired, to be copied onto the first page of output for the cycle.
Card 2, 1–72	Continuation of title.

Although these cards may be left blank, they must both be present.

Parameter Card (3)

This card tells the main program the basic parameters of the problem and options which are desired.

Columns	*Contents*
1– 4	Number of observations (the Data Cards or sets of Data Cards) to be processed.
5– 6	Number of variables to be processed ($\leq$ 40).
7– 8	Number of tables to be constructed ($\leq$ 72).

continued

Columns	*Contents*
	Note: In Col. 9-22, 00 or blank indicates that the option is not desired; 01 that the option is desired.
9–10	Variable identification option: This permits the labeling of each table with an alphanumeric description of each variable involved in the cross-classification. If this feature is used, include the Variable Identification Cards.
11–12	Recoding option: This allows the variables to be recoded in the machine by adding, subtracting or dividing by a constant c, where c is any positive integer less than 100. It also allows one to five digit data to be regrouped into intervals or a specified size. If the feature is desired, include the Recoding Control Cards.
13–14	Tape input option: This permits use of an input tape instead of card input. If the tape is used, a DD card must be defined in the job control for data set reference number 2. The blocking factor, record size, and labeling information contained on the DD card will depend on how the individual tape was created. (See Ref. 1)
15–16	Rewind option: The input data tape will rewind after each execution of the program. This allows processing more than 40 variables or more than 72 tables in one computer run with the same data. This option should only be used when an 01 has been punched in Col. 13–14.
17–18	Chi square option: The chi square value and the contingency coefficient (C) will be computed for each bivariate distribution in every table. The zero row and zero column frequencies are *omitted* from these calculations.$^{\alpha}$
19–20	Column and row percentages: The tables containing the percentages by row and by column will be printed. The zero row and column frequencies will be eliminated from these percentages.$^{\alpha}$ Percentages which cannot be computed because the row or column total is zero will be printed as 0.0.
21–22	Correlation coefficient option: Several correlation coefficients will be computed, depending on the number of rows and columns.

continued

Columns *Contents*

Size of Table	*Correlation Coefficients*
n x n	Kendall Tau-B, Goodman-Kruskal Gamma, Somers D_{yx} and D_{xy}.
n x m	Kendall Tau-C, Goodman-Kruskal Gamma, Somers D_{yx} and D_{xy}.

The size of table refers to the number of rows and columns after eliminating the zero categories.[α] The correlation coefficients which can not be computed because of division by zero are printed as ******.

[α] If the zero category for any variable is desired in the calculation, the variable can be recoded by adding 1 to every category.

Format Cards (4–6)

The format of the variables on the Data Cards should be punched in Col. 1–72 of Card 4 and continued in Col. 1–72 of Cards 5 and 6, if necessary. The computer will regard the first variable it reads as variable #1, the next one as variable #2, and so on, in consecutive order reading the card from left to right. The user must refer to the variables by this same number when he tells the machine which tables to build. All three cards must appear even if some of them are blank. *CAUTION:* The variables must be in I-format.

Maxima Card (7)

The maximum values of the variables should be punched in the 40 two-digit fields corresponding to the variables read in: the first field (Col. 1–2) corresponds to variable #1, etc. If the variable has been recoded, either the maximum value after recoding (for codes 01–03) or the number of categories after regrouping (for codes 04–07) should be used. (See below). In general, the values given should not be larger than the true maximum as the results may be less accurate and the job time longer.

Note: The variables which are to appear along the columns must have a maximum value of 20.

continued

Tables Cards (8–11)

These cards designate the variables to be used in building the tables. The user identifies these variables by the order in which they were read in. Each two column field on Card 8 names the variable which will appear along the columns of a certain table. (Note that the maximum value of these variables must be 20 as no table has more than 20 columns.) The corresponding two column field in Card 9 names the variable to appear in the rows of that table. If the user wants to cross-classify the pair of variables while holding a third variable constant, he should punch the number of the third variable into the corresponding columns of Card 10. A fourth variable can be introduced into the table the same way with the use of Card 11. The first field (Col. 1–2) for each of these cards corresponds to the first table the computer will construct; the second field (Col. 3–4) corresponds to the second table, and so on.

When a third variable is used in a table, a separate sub-table is prepared for each category of the third variable. Understand that one sub-table will be generated for *each* category of a third variable. It is possible to generate up to 99 such sub-tables for each table. The introduction of third variables into the analysis should be done with attention to the number of sub-tables to be generated in the process. Try to use as third variables those with few categories, using those with more categories as first and second variables. Be even more careful when introducing a fourth variable, for the number of sub-tables produced will be equal to the number of categories for the third variable multiplied by the number of categories for the fourth.

If more than 40 tables are to be created, Cards 8–11 must each be continued on a second card. The variables for tables 1–40 should be punched on Cards 8, 9, 10, 11. The variables for tables 41–72 should be punched on Cards 8a, 9a, 10a, 11a. These cards must be in the following order: 8, 8a, 9, 9a, 10, 10a, 11, 11a.

Recoding Control Cards

If the recoding option is used, two cards must be read in immediately after Card 11 of the Instruction Cards. These cards contain 40 two-column fields. The first such field, in Col. 1–2, corresponds to the variable #1, and so on, in ascending order to variable #40 in Col. 79–80. The first of the two cards contains the code for recoding. If the variable is not to be recoded, the corresponding field should be left blank. The codes are as follows:

continued

Code	*Value*
01	Recode by adding the constant c to the value.
02	Recode by subtracting the constant c from the value.
03	Recode by dividing the constant c into the value.
04	Recode by regrouping the data into specified intervals.
05	Recode by first regrouping the data into specified intervals and then adding c to the interval code.
06	Recode by first regrouping the data into specified intervals and then subtracting c from the interval code.
07	Recode by first regrouping the data into specified intervals and then dividing c into the interval code.

Note: If options 04–07 are exercised, the intervals are assigned code numbers by the machine. The first interval is assigned one, the second interval two and so on. These interval codes thus become the values of the variables that are regrouped, and are themselves recoded if options 05–07 are exercised. Note that the zero category is omitted from the calculations.

The constants to be used for addition, subtraction, or division should be punched in the appropriate 2 digit field on the second Recoding Control Card. Any value which becomes negative when it is recoded is set to zero.

If regrouping is used (codes 04–07), a Regrouping Card for each variable recoded in this manner must follow the two control cards. These cards contain 16 five-column fields for the categories into which the variable will be regrouped. The first field, Col. 1–5, should contain the upper limit of the first interval into which the data are to be regrouped. The second field should contain the upper limit of the second interval, etc. The fields on Recoding Card #2 should be left blank if code 04 is used. However, this second card must be present, even if regrouping alone is being done and the card is blank. The regrouping will be done before addition, subtraction, or division.

Variable Identification Cards

If the variable identification option is used, a description card for each variable must immediately precede the Data Cards. An alphanumeric description of the variable should be punched into Col. 1–72 of the Variable Identification Cards. The cards must be arranged in the same sequence as the variables are arranged on the Data Cards.

continued

Raw Data

The raw data are assumed to be positive integers of value less than 100. If the data are larger than 100, they must be recoded. The data must be in I format but may be arranged in any manner on one or more cards since the user provides the format cards, which are Cards 4, 5, and 6 on the Instruction Cards described above.

Additional Notes

The program is written in full FORTRAN IV and has been used under OS. It requires about 60K bytes of core storage on an IBM S/360 and uses a work area which may be defined either on disk or tape. This work area uses data set reference number 04. If using NUCROS from STATLIB (a library of statistical programs at the University of Iowa) this work area is predefined.

On an IBM 360/65 under OS and HASP, one cycle of 200 data sets with 40 variables and 40 tables took 2.5 minutes.

Caution

This program produces a large amount of output, so local users will need to specify the number of lines to be printed on their JOB Card. Forty tables require at least 120 pages or 6000 lines of print.

REFERENCES

1. IBM S/360 Job Control Language Manual C28-6539 (IBM Corporation, White Plains, New York).

SAMPLE INPUT

```
      TEST CASE FOR NUCROS WITHOUT RECODING OR IDENTIFICATION
      LEWIS     JULY  1967
  202008          010101
(2X,10I1,2X,2(5I1,2X))


0305010204020302020102040301020604010203
0205071213161720
040106080911151 9
    03  18    14
    10
A 11101031212431241123    00100      111001
B 03024112101220064010    12141 1200310112
C 25013210210130053111    2325017500221310
D 32022030210330212021    34330 9000007312
E 24121201102020221110    42651 3200205011
F 11114131101431230003    51521  600325102
G 02022132210120064121    61420  200112000
H 05103032012130255023    70200   10017002
I 23111211101431011111    80601   35003001
J 24013131102321042001    01110  125120101
K 34124211210031233113    11011   50305002
L 12022010210230214002    20000     215101
M 13011200212420254120    46520 4500309400
N 35005112101430124012    62841 7500106212
O 21122032102330061101    8353012000116210
P 02124230101231010023    74150 8000023312
Q 02111011010121262112    53121 6000213100
R 33112130010221253000    31931 1100303002
S 24023010101331243121    2543118000201010
T 14104111102431012033    41220  125117302
      TEST CASE FOR NUCROS WITH RECODING, WITHOUT IDENTIFICATION
     LEWIS    JULY  1967
  201106  01      01
(14X,I1,10X,I1,I2,2I1,I5,I1,I2,3I1)


01090803010403050401 02
0203040608 09
0711010510 01
    10

02010504  04  03
020101        05
    4   13   18   30   45   60  100
    2    4    6
 4000 75001200020000
A 11101031212431241123    00100      111001
B 03024112101220064010    12141 1200310112
C 25013210210130053111    2325017500221310
D 32022030210330212021    34330 9000007312
E 24121201102020221110    42651 3200205011
F 11114131101431230003    51521  600325102
G 02022132210120064121    61420  200112000
H 05103032012130255023    70200   10017002
I 23111211101431011111    80601   35003001
J 24013131102321042001    01110  125120101
K 34124211210031233113    11011   50305002
L 12022010210230214002    20000     215101
M 13011200212420254120    46520 4500309400
N 35005112101430124012    62841 7500106212
O 21122032102330061101    8353012000116210
P 02124230101231010023    74150 8000023312
Q 02111011010121262112    53121 6000213100
R 33112130010221253000    31931 1100303002
S 24023010101331243121    2543118000201010
T 14104111102431012033    41220  125117302
/*
```

SAMPLE OUTPUT

```
     TEST CASE FOR NUCROS WITHOUT RECODING OR IDENTIFICATION
     LEWIS    JULY  1967

TOTAL CASES USED =   20     CYCLE NO.   1 OF   2

VARIABLE   1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20
MAX.VALUE  3  5  1  2  4  2  3  2  2  1  2  4  3  1  2  6  4  1  2  3

         TABLE NO.       VARIABLES        MAXIMUM VALUES
             1            2   4            5   2
             2            5   1            4   3
             3            7   6   3  10    3   2   1   1
             4           12   8            4   2
             5           13   9  18        3   2   1
             6           16  11            6   2
             7           17  15            4   2
             8           20  19  14        3   2   1

CARD 1       1    1    1    0    1    0    3    1    2    1    3    1    2    4    1    3    0    0
        0    0

CARD 2       0    3    0    2    4    1    1    2    1    0    2    0    0    6    4    0    0    0
        0    1

CARD 3       2    5    0    1    3    2    1    0    2    1    3    0    0    5    3    1    0    0
        0    2




TABLE NO.=  1        TABLE SIZE=  6 BY   3
                     VARIABLE NO.  2 VS. NO.  4

             TOT        0    1    2    3    4    5

      0       4         0    1    0    0    1    2
      1       7         0    1    1    3    1    1
      2       9         0    1    4    1    3    0

    TOTAL    20         0    3    5    4    5    3

   CHI SQUARE =   4.622   C=0.473     TAU-C = -0.156   GAMMA = -0.192   DXY = -0.159   DYX = -0.103

   PERCENTS BY COLUMN FROM THE ABOVE MATRIX

             TOT             1    2    3    4    5

      1       7           50.0 20.0 75.0 25.0100.0
      2       9           50.0 80.0 25.0 75.0  0.0

    TOTAL    16              2    5    4    4    1

   PERCENTS BY ROW FROM THE ABOVE MATRIX

             TOT             1    2    3    4    5

      1       7           14.3 14.3 42.9 14.3 14.3
      2       9           11.1 44.4 11.1 33.3  0.0

    TOTAL    16              2    5    4    4    1

        .                         .
        .                         .
        .                         .
```

continued

```
TABLE NO.=  7        TABLE SIZE=  5 BY   3
                     VARIABLE NO. 17 VS. NO. 15

                TOT       0    1    2    3    4

         0       2        2    0    0    0    0
         1       4        4    0    0    0    0
         2       4        4    0    0    0    0

      TOTAL     10       10    0    0    0    0

TABLE NO.=  8        TABLE SIZE=  4 BY   3
                     VARIABLE NO. 20 VS. NO. 19  WITH CONSTANT VARIABLES  14
                     3RD DIMENSION =   0

                TOT       0    1    2    3

         0       0        0    0    0    0
         1       0        0    0    0    0
         2       0        0    0    0    0

      TOTAL      0        0    0    0    0


TABLE NO.=  8        TABLE SIZE=  4 BY   3
                     VARIABLE NO. 20 VS. NO. 19  WITH CONSTANT VARIABLES  14
                     3RD DIMENSION =   1

                TOT       0    1    2    3

         0       2        0    0    1    1
         1       0        0    0    0    0
         2       0        0    0    0    0

      TOTAL      2        0    0    1    1




      TEST CASE FOR NUCROS WITH RECODING, WITHOUT IDENTIFICATION
     LEWIS    JULY  1967

TOTAL CASES USED =   20     CYCLE NO.   2 OF   2

VARIABLE     1  2  3  4  5  6  7  8  9 10 11
MAX.VALUE    1  9  8  3  1  4  3  5  4  1  2

          TABLE NO.       VARIABLES        MAXIMUM VALUES
               1             2  7             9  3
               2             3 11             8  2
               3             4  1  10         3  1   1
               4             6  5             4  1
               5             8 10             5  1
               6             9  1             4  1
```

continued

```
RECODING OPTION CALLED FOR

    VAR.NO. CODE       CONSTANT      RECODED CATEGORIES (UPPER LIMITS)
       1      2           2
       2      1           1
       3      5           1           4    13    18    30    45    60   100     0     0     0     0     0     0     0     0     0
       4      4                       2     4     6     0     0     0     0     0     0     0     0     0     0     0     0     0
       6      4                    4000  7500 12000 20000     0     0     0     0     0     0     0     0     0     0     0     0
       8      3           5

CARD 1          3    0    1    0    0     0    1   11    0    0    1

CARD 2          2    1   21    4    1  1200    3   10    1    1    2

CARD 3          3    2   32    5    017500     2   21    3    1    0


TABLE NO.=  1       TABLE SIZE= 10 BY   4
                    VARIABLE NO.  2 VS. NO.  7

               TOT        0     1     2     3     4     5     6     7     8     9

        0        4        0     0     0     0     1     0     0     0     2     1
        1        6        0     2     0     0     0     1     0     2     0     1
        2        5        0     0     0     3     0     1     1     0     0     0
        3        5,       0     0     2     0     1     1     1     0     0     0

    TOTAL       20        0     2     2     3     2     3     2     2     2     2

    PERCENTS BY COLUMN FROM THE ABOVE MATRIX

               TOT              1     2     3     4     5     6     7     8     9

        1        6          100.0   0.0   0.0   0.0  33.3   0.0100.0   0.0100.0
        2        5            0.0   0.0100.0    0.0  33.3  50.0   0.0   0.0   0.0
        3        5            0.0100.0    0.0100.0   33.3  50.0   0.0   0.0   0.0

    TOTAL       16              2     2     3     1     3     2     2     0     1

    PERCENTS BY ROW FROM THE ABOVE MATRIX

               TOT              1     2     3     4     5     6     7     8     9

        1        6           33.3   0.0   0.0   0.0  16.7   0.0  33.3   0.0  16.7
        2        5            0.0   0.0  60.0   0.0  20.0  20.0   0.0   0.0   0.0
        3        5            0.0  40.0   0.0  20.0  20.0  20.0   0.0   0.0   0.0

    TOTAL       16              2     2     3     1     3     2     2     0     1

               .                                  .
               .                                  .
               .                                  .
```

continued

```
TABLE NO.=  6        TABLE SIZE=  5 BY   2
                     VARIABLE NO.   9 VS. NO.  1

                 TOT          0    1    2    3    4

           0       7          3    3    0    0    1
           1      13          5    2    2    4    0

       TOTAL      20          8    5    2    4    1



      PERCENTS BY COLUMN FROM THE ABOVE MATRIX

                TOT              1    2    3    4

           1      8          100.0100.0100.0  0.0

       TOTAL      8              2    2    4    0

      PERCENTS BY ROW FROM THE ABOVE MATRIX

                TOT              1    2    3    4

           1      8           25.0 25.0 50.0  0.0

       TOTAL      8              2    2    4    0

2 CYCLES COMPLETED - THIS JOB DONE
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $0.81.

Charge to user = computer costs + postage + network overhead
= $0.81 + postage + network overhead

CONTENTS—NUCROS

pages
1 Identification & Abstract
3– 8 User Instructions
9–13 I/O
15 Cost—Contents

DESCRIPTIVE TITLE	Data Survey and Normality Test
CALLING NAME	NORMSURV
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Carol Hopkins Yale University Computer Center Revised by Louise Levine The University of Iowa Computer Center
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Louise R. Levine, Program Librarian, University Computer Center, The Univ. of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

NORMSURV allows the user to test the hypothesis that his data are normally distributed, using a chi-square test. Missing data are detected as either minus zero or blank in the input data. Values of zero are considered to be valid and are used in all calculations. Means, standard deviations and a list of outliers (see below) will be indicated on the output. Transformations are allowed to generate new variables or to change any variable before the analysis is done. A median and test of skewness are also calculated.

The data are scanned variable by variable for missing data. Means and standard deviations are calculated using only the number of non-missing observations. The data are then standardized (mean = 0, variance = 1), checked for extremes, and placed in a 10-cell frequencies distribution. The limits of the cells are the deciles of the normal distribution of standardized data. The estimate of chi-square has seven (7) degrees of freedom because three degrees are lost in estimating μ by $\bar{x}$, σ by s, and forcing $\sum_{i=1}^{10} f_e = k_n$ where f_e is the expected frequency in each cell.

continued

Output includes the following:

—list of transformation cards,
—list of data before and after transformation (optional),
—list of missing observations,
—number of non-missing observations, sample mean and standard deviation
—list of observations whose standard scores are greater in absolute value than an extreme value of the user's choice (outliers),
—a graph of the cell frequencies
—an estimated chi-square statistic.

USER INSTRUCTIONS

Input Deck

The input deck consists of the following sets of cards in the order described.

Title Card

Columns	*Contents*
1–72	Any alphanumeric information.

Problem Card

Columns	*Contents*
1– 2	Problem number.
3– 4	The number of variables, N, ($\leq$ 50).
5– 8	The number of observations, K, [$K*(N+1) \leq 18,000$].
9–10	The number of Transformation Cards.
11–12	The number of Selection Cards.
13	5: card input Tape number for special tape unit.
15–19	The extremes limit. The number placed in these columns represents the number of standard deviations from the mean that are acceptable (the decimal point must be punched).
20	1: A listing of the data before and after transformation is desired. 0: No listing of the data is desired
21–22	The number of variables added during transformation.
23	The number of Variable Format Cards for input ($\leq$ 3).

Variable Format Cards

A maximum of 3 cards are allowed using only Col. 1–72. An E or F type format is needed to read in one observation of N variables.

Variable Name Cards

Columns	*Contents*
1–12	Name of first variable.
13–24	Name of second variable.
⋮	⋮
61–72	And so on, 12 columns to a name, 6 names to a card, until all of the variables are named.

continued

Observation Cards

Punch according to the format on the Variable Format Cards.

Transformation Cards

Transformations are performed in the order in which they occur and are retained throughout the rest of the analysis. However, a transformation may be specified for each selection by using Selection Cards. Then it holds only for that selection.

Columns	*Contents*
1– 3	The variable number assigned to X'. This may be the same as the present number.
4– 5	Transformation Code.
6–10	The variable number of X_A.
11–16	The variable number of X_B for transformation codes 11–14 or the constant C for codes 1–10. The constant should be right adjusted or the decimal point should be punched.

Transformation Codes are as follows.

Notation: * indicates multiplication
** indicates raised to a power

Code	*Transformation*	*Restriction*
01	$X' = X_A+C$	
02	$X' = X_A*C$	
03	$X' = X_A**C$	$X_A \geq 0$
04	$X' = \sqrt{X_A-C}$	$X_A \geq C$
05	$X' = EXP(X_A+C)$	$X_A+C<88.028$
06	$X'' = ASIN(X_A+C)$	$-1 \leq X_A+C \leq 1$
07	$X' = ATAN(X_A+C)$	
08	$X' = 1/(X_A+C)$	$X_A + C \neq 0$
09	$X' = LOG_{10}(X_A+C)$	$X_A+C>0$
10	$X' = LOG_E(X_A+C)$	$X_A+C>0$

continued

Code	*Transformation*	*Restriction*
11	$X' = X_A + X_B$	
12	$X' = X_A - X_B$	
13	$X' = X_A * X_B$	
14	$X' = X_A / X_B$	$X_B \neq 0$

The violation of any of the above restrictions during the transgeneration phase of the program causes execution to stop. Any violation during the transformation phase causes a transfer to the next selection. In either case, the offending observations are listed.

Selection Cards

There may be from one to eight selections on a card. The number in Col. 11–12 of the Problem Card is the number of *cards*, not the number of selections. The specification of each variable takes nine columns as follows.

Columns	*Contents*
1– 2	Number of variable being analyzed in this selection.
3– 4	The Transformation Code (transformation codes 01 through 10 are allowed where X' and X_A are the variable specified in the two preceding columns, *or code 00 meaning no transformation*).
5– 9	The constant C for transformation codes 01 through 10. Leave blank for other codes.

Repeat the above information in Col. 10–18 for the second selection on the card, in Col. 19–27 for the third, etc.

END Card

Columns	*Contents*
1– 3	END

/* Card

Columns	*Contents*
1– 2	/*

continued

000 0117

The END Card and the /* card follow the complete run. If more than one problem is to be run, repeat Title Card through Selection Cards before the END Card and the /* Card.

Note: If a special tape is used for input, the unit must be defined using DD cards. NORMSURV requires 128K bytes for execution and uses an OS Assembler Language program, BLNK.

SAMPLE INPUT

```
TESTING DATA SCREENING AND NORMALITY TEST PROGRAM
013101000002501.0001001
(F3.0,1X,2F2.0,F3.0,2F2.0,F3.0,8F1.0,4F2.0,12F1.0)
CASENO     PATIENT NO. HOSPITAL NO.     SEX      RACE       DATE
       YEARNO. OF ADMITDIAGNOSIS     HEIGHT    WEIGHT     HGT/WT
     FEVER1       FEVER2     FEVER3  BLDPR1    BLDPR2     BLDPR3
       URN1         URN2       URN3    EKG1      EKG2       EKG3
    RESULT1      RESULT2    RESULT3   DRUG1     DRUG2      DRUG3
POST EXAM
529 12 6 54 4 4 3336253231 1144312433223101243121516
530 12 6 24 5 6 63212233233 155212253323 91353132612
533 13 6 14 610 532          43 12454422    631216
534X14 5 74 5 4 632683121452267412243524121843222623
536 14 6 24 611 73398311314226324  3531     1 121536
537 14 4 62 8   91215  231 2 2111  52 2320206 21123
538X12 6 54 1 7 6333222  1 2 53 1237322320206 222324
539 14 6 34 913 633584112332243312343424 91444122444
540 14 6 33 3 4 232453111131 2313  71 1     7 342625
542 14 6 13 1 1 134323111131231311 5352220  64132624
501 14 6 741017 523423122122124211 6342218  63132525
502 14 6 34 710 8336231111  142111 53422 6  53121623
503 14 6 34 1   144525321422164111 42425 72044122612
507 14 6 34 9 7 9244853  1 2162111 43523 1 534321545
508 13 6 74 4 2 2248853    1164 11 33222 3  3312214
509 11 6 14 4 4 423323213132227111 43224 21733121642
510X14 6 34 4 5 5223241323 1172132333325 41233241232
511 11 6 14 1    121         11111 73121    643117
513 13 6 24 6 9 423523112252115111 45528 31834121131
514 14 6 44 5 6 423825333132162111 26322 2  33122432
471 14 5 731118 823683312352162 12333424 51034132621
472 13 5 44 912 42232        62311 42325102044122636
473 14 4 63 6 8 61288533313117114  1321     1 121642
478X11 6 54 1 1 22352423 321143111 7112520206 222623
479 13 6 34 6 7 533323111132253511 53523141853322623
480 13 6 54 5 5 63332 3  1  131111 73 2220  6312263
481 13 6 14 1    334231111 1125111 63523202064122633
482 14 6 34 7 8 533623131132262212243224 01833122533
486 13 3 51 5 5 523725311 3214513  42223192043122632
487 11 6 14 1    1452532212211113  56525202064132611
417 14 6 54 916 623323332152235211 36424 6 944221646
418 14 6 54 8 4 7236231  1 2172111 71128     6 321141
419 14 6 64 6 9 523623332132172112236523 7 944122634
420 13 6162 7 8 8421553    1 15112351 2317205332161
421 12 4 72 6 6 53332431323213313  2431     1 322533
422X14 6 14 1 1 213123233132232111 7211     1 231111
423 13 6 14 4 5  3312        1313  62323    63121612
427 14 5 23 8 9 724581      143111 44423101444122531
429 11 6 44 5 8 5235231121 2241111 43329 32033122611
430 12 6 34 5 5 4331231112 2 41212353323152053  2623
431X12 2 24 612 4231233221 1123112 32324 4 733221411
432 14 6 44 8 9 53482 2133 2 41512245423 81344122646
433 14 6 34 7 6 933225322121275111 36222 3  34122626
434 13 6 34 910 62212       231111 35527 41434212623
434X13 6 74 910 72212       231111 35527 41434222624
437 14 6 6410 2 83378411133214213  54424192064212433
438 14 6 34 911 633925312132166111 23423 4 533122612
439 14 4 42 911 8332241114 1135212435324 4 734141114
440 14 4 11 3 2 324223111152251 3  34323 1 53 131221
441 11 6 54      421         2111224432315194 122643
442 14 6 14 5 8 6225211  3   54212353221    63122621
443 13 5 11 1    3212311115114613  34221    2 122521
447 13 6 14 4 4 4331253331 2146112313223 0 133121321
450 14 6 54 5 7 3237232111 2156111 43325 92043122623
451 13 6 64 5 7 42375       154111 42321    23312621
452 14 6 94 4 4 533285332451237111 13222 1  33122611
453 14 5 63 7 5 623753111452163132353223202063121525
454 13 6 54 2   3222232   11 311   7111     7 122641
454X12 6 24 5   4226232  1  27523  6432320206 222642
457 14 6 641313 72358321 22216311236122     5 121224
458 14 6 24 610 523583311132161612334324 5 933121625
459 14 6 24 912 63338123 2 113611235442211  53132513
460 14 6 52 2   223 231    2 23211 7612220  631222 4
```

continued

```
461 14 6 84      332         31111 23323 3 533122523
462X14 6 34 6 9 6336 53 23 2154211 33325 2 734222722
463 14 6 24 510 533623111152175111 33323 2 933121631
464 13 6 71 1    222253211321233              122636
466 14 6 24 7 7 422753222332273312342527 41534132643
467 14 5 54 8 8 734825313332163112333324 3 734121421
469 13 6 24 2 2 234525312122154112253423202064322621
488 12 6 53 7 9 333324111232122212244324151843121612
489 14 6 44 4 6 4236231131 2164111 5322220  63122646
491 13 6 74 8 7 633654312332256112353324132053122622
492 12 6 44 2 1 333524231332223112353325 92053122611
493 13 6 14 5 2 42382411331216711235242220  64122623
496 13 6 84 5 6 323325312122154112322325 3 633131521
497 14 6 24 5 8 523622221232276512356424122054132427
498 14 6 24 1 2 223123332132145 11 33223 3 433132631
499 13 6 24 5 7 533825311132143312234521    24122625
500 12 6 54 3 2 5337231113 215613  51 25202063241211
516X13 6 44 8 5 633285331132176112214221    2 231334
517 13 6 24 711 73332433345215324  45421    23121625
518 13 6 34 916 923223223152127111 43222 9  44122633
520 14 6 33 913 7345831221 2221411 43525 81644122624
521 13 6 24 910 632323112132147212343529 72044122623
522 13 6 14 5 6 624653232132167 11 36324 21133122621
524 12 5 31 2 1 2233212  15222344  44424 7124 122524
526 12 6 51 3 5 323223311121243 2  5442320206 122611
527 13 6 34 2 3 123645311132161111 33422 3  34122625
528 14 4 12 4 5 52342533313114712  2321     1 122531
543 14 6 44 2 3 33272411332216413  6311520206 132611
544X14 6 74 814 933984131 2217623  6122420206 221245
545 13 6 13 810 624425312132143112233323 0 234122613
545A13 6 54 8 5 723584112 32256112243325 91643121222
546 12 5 14 3 4 23412412123213114  4541     1   2611
547 12 6 31 9 8 6331531111 1131311 7442220  6 122624
549 14 6 24 5 6 433685331122167111 33321    23  2624
550 14 6 34 5 9 632823111132166111 33423 1 634152624
551 14 6 54 914 9336833111 2162112224423 1 434132432
552 13 6 24 5 6 233425313122142 12233322 3  33131541
0100     0201 1.000302 2.000400      0500     0600     0700      0800
09041.000
END
//
```

SAMPLE OUTPUT

```
PROBLEM NO.  1             TESTING DATA SCREENING AND NORMALITY TEST PROGRAM

NO. OF VARIABLES = 31       NO. OF OBS. =  100       NO. OF TRANSGEN. CARDS =  0       NO. OF SELECTION CARDS =  2       NTAPE = 5

      SELECTION   1                VARIABLE NO.  1               CASENO                    TRANSFORMATION NO.  0

      DATA BEFORE TRANSFORMATION
        OBS. NO.

         1-   8    5.2900E 02  5.3000E 02  5.3300E 02  5.3400E 02  5.3600E 02  5.3700E 02  5.3800E 02  5.3900E 02
         9-  16    5.4000E 02  5.4200E 02  5.0100E 02  5.0200E 02  5.0300E 02  5.0700E 02  5.0800E 02  5.0900E 02
        17-  24    5.1000E 02  5.1100E 02  5.1300E 02  5.1400E 02  4.7100E 02  4.7200E 02  4.7300E 02  4.7800E 02
        25-  32    4.7900E 02  4.8000E 02  4.8100E 02  4.8200E 02  4.8600E 02  4.8700E 02  4.1700E 02  4.1800E 02
        33-  40    4.1900E 02  4.2000E 02  4.2100E 02  4.2200E 02  4.2300E 02  4.2700E 02  4.2900E 02  4.3000E 02
        41-  48    4.3100E 02  4.3200E 02  4.3300E 02  4.3400E 02  4.3400E 02  4.3700E 02  4.3800E 02  4.3900E 02
        49-  56    4.4000E 02  4.4100E 02  4.4200E 02  4.4300E 02  4.4700E 02  4.5000E 02  4.5100E 02  4.5200E 02
        57-  64    4.5300E 02  4.5400E 02  4.5400E 02  4.5700E 02  4.5800E 02  4.5900E 02  4.6000E 02  4.6100E 02
        65-  72    4.6200E 02  4.6300E 02  4.6400E 02  4.6600E 02  4.6700E 02  4.6900E 02  4.8800E 02  4.8900E 02
        73-  80    4.9100E 02  4.9200E 02  4.9300E 02  4.9600E 02  4.9700E 02  4.9800E 02  4.9900E 02  5.0000E 02
        81-  88    5.1600E 02  5.1700E 02  5.1800E 02  5.2000E 02  5.2100E 02  5.2200E 02  5.2400E 02  5.2600E 02
        89-  96    5.2700E 02  5.2800E 02  5.4300E 02  5.4400E 02  5.4500E 02  5.4500E 02  5.4600E 02  5.4700E 02
        97- 100    5.4900E 02  5.5000E 02  5.5100E 02  5.5200E 02

       THERE IS NO MISSING DATA

       SAMPLE SIZE =  100
       MEAN =          485.7598
       STD. DEV. =      40.9563
       MEDIAN =          486.5000
       SKEW =          -0.0542

       EXTREMES GREATER THAN  1.000 OR LESS THAN  -1.000. . .
        OBS. NO.              TRANSFORMED VALUE     STANDARDIZED VALUE


           1                    5.29000E 02             1.0558
           2                    5.30000E 02             1.0802
           3                    5.33000E 02             1.1534
           4                    5.34000E 02             1.1778
           5                    5.36000E 02             1.2267
           6                    5.37000E 02             1.2511
           7                    5.38000E 02             1.2755
           8                    5.39000E 02             1.2999
           9                    5.40000E 02             1.3243
          10                    5.42000E 02             1.3732
          31                    4.17000E 02            -1.6789
           .
           .                        .                      .

           .                        .                      .
          52                    4.43000E 02            -1.0440
          89                    5.27000E 02             1.0069
          90                    5.28000E 02             1.0313
          91                    5.43000E 02             1.3976
          92                    5.44000E 02             1.4220
          93                    5.45000E 02             1.4464
          94                    5.45000E 02             1.4464
          95                    5.46000E 02             1.4708
          96                    5.47000E 02             1.4953
          97                    5.49000E 02             1.5441
          98                    5.50000E 02             1.5685
          99                    5.51000E 02             1.5929
         100                    5.52000E 02             1.6173
```

continued

```
PROBLEM NO.  1               TESTING DATA SCREENING AND NORMALITY TEST PROGRAM
SELECTION   1                VARIABLE NC.  1             CASENO                     TRANSFORMATION NO.  0
FREQUENCY

   38 -
   36 -
   34 -
   32 -
   30 -
   28 -
   26 -                                                                                          X
                           X
   24 -
   22 -
   20 -
   18 -
   16 -
   14 -
   12 -                              X
   10 -                                                                                X
    8 -                                                            X         X
    6 -                                        X
                                                         X
    4 -
    2 -
    0 -          X                                                                                         X
DECILES     I         I         I         I         I         I         I         I         I         I         I
          -INF    -1.282    -0.842    -0.524    -0.253        0      0.253     0.524     0.842     1.282     +INF

      N          0        25        12         6         5         8         8        10        26         0    TOTAL =  100

      +          0.0      15.0       2.0       0.0       0.0       0.0       0.0       0.0      16.0       0.0

      -        -10.0       0.0       0.0      -4.0      -5.0      -2.0      -2.0       0.0       0.0     -10.0

   EXPECTED FREQUENCY =  10.0

   CHI-SQUARE =    73.4000         D. F. = 7
```

continued

```
PROBLEM NO.  1             TESTING DATA SCREENING AND NORMALITY TEST PROGRAM

SELECTION   3              VARIABLE NO.  3              HOSPITAL NO.              TRANSFORMATION NO.  2

DATA BEFORE TRANSFORMATION
  OBS. NO.


    1-   8   6.0000E 00  6.0000E 00  6.0000E 00  5.0000E 00  6.0000E 00  4.0000E 00  6.0000E 00  6.0000E 00
    9-  16   6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00
   17-  24   6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  5.0000E 00  5.0000E 00  4.0000E 00  6.0000E 00
   25-  32   6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  3.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00
   33-  40   6.0000E 00  6.0000E 00  4.0000E 00  6.0000E 00  6.0000E 00  5.0000E 00  6.0000E 00  6.0000E 00
   41-  48   2.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  4.0000E 00
   49-  56   4.0000E 00  6.0000E 00  6.0000E 00  5.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00
   57-  64   5.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00
   65-  72   6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  5.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00
   73-  80   6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00
   81-  88   6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  5.0000E 00  6.0000E 00
   89-  96   6.0000E 00  4.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00  5.0000E 00  6.0000E 00
   97- 100   6.0000E 00  6.0000E 00  6.0000E 00  6.0000E 00

DATA AFTER  TRANSFORMATION
  OBS. NO.


    1-   8   1.2000E 01  1.2000E 01  1.2000E 01  1.0000E 01  1.2000E 01  8.0000E 00  1.2000E 01  1.2000E 01
    9-  16   1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01
   17-  24   1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.0000E 01  1.0000E 01  8.0000E 00  1.2000E 01
   25-  32   1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  6.0000E 00  1.2000E 01  1.2000E 01  1.2000E 01
   33-  40   1.2000E 01  1.2000E 01  8.0000E 00  1.2000E 01  1.2000E 01  1.0000E 01  1.2000E 01  1.2000E 01
   41-  48   4.0000E 00  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  8.0000E 00
   49-  56   8.0000E 00  1.2000E 01  1.2000E 01  1.0000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01
   57-  64   1.0000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01
   65-  72   1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.0000E 01  1.2000E 01  1.2000E 01  1.2000E 01
   73-  80   1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01
   81-  88   1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.0000E 01  1.2000E 01
   89-  96   1.2000E 01  8.0000E 00  1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01  1.0000E 01  1.2000E 01
   97- 100   1.2000E 01  1.2000E 01  1.2000E 01  1.2000E 01


THERE IS NO MISSING DATA

SAMPLE SIZE =  100
MEAN =           11.4400
STD. DEV. =       1.4236
MEDIAN =           12.0000
SKEW =          -1.1801

EXTREMES GREATER THAN  1.000 OR LESS THAN  -1.000. . .
 OBS. NO.                  TRANSFORMED VALUE     STANDARDIZED VALUE


      4                      1.00000E 01            -1.0115
      6                      8.00000E 00            -2.4164
     21                      1.00000E 01            -1.0115
     22                      1.00000E 01            -1.0115
     23                      8.00000E 00            -2.4164
     29                      6.00000E 00            -3.8213
     35                      8.00000E 00            -2.4164
     38                      1.00000E 01            -1.0115
     41                      4.00000E 00            -5.2261
     48                      8.00000E 00            -2.4164
     49                      8.00000E 00            -2.4164
     52                      1.00000E 01            -1.0115
     57                      1.00000E 01            -1.0115
     69                      1.00000E 01            -1.0115
     87                      1.00000E 01            -1.0115
     90                      8.00000E 00            -2.4164
     95                      1.00000E 01            -1.0115
```

continued

```
PROBLEM NO.  1               TESTING DATA SCREENING AND NORMALITY TEST PRCGRAM
SELECTION   3                VARIABLE NC.  3              HOSPITAL NC.               TRANSFORMATION NO.  2
FREQUENCY

   95 -
   90 -
   85 -
                                                                            X
   80 -
   75 -
   70 -
   65 -
   60 -
   55 -
   50 -
   45 -
   40 -
   35 -
   30 -
   25 -
   20 -
                          X
   15 -
   10 -
    5 -
    0 -         X                   X         X         X         X                   X         X         X
DECILES    I         I         I         I         I         I         I         I         I         I         I
         -INF     -1.282    -0.842    -0.524    -0.253       0       0.253     0.524     0.842     1.282     +INF
      N          0        17         0         0         0         0        83         0         0         0    TOTAL =  100
      +         0.0       7.0       0.0       0.0       0.0       0.0      73.0       0.0       0.0       0.0
      -        -10.0       0.0     -10.0     -10.0     -10.0     -10.0       0.0     -10.0     -10.0     -10.0

   EXPECTED FREQUENCY =  10.0
   CHI-SQUARE =   617.7998        D. F. = 7
```

continued

```
              .                                      .                                      .
              .                                      .                                      .
              .                                      .                                      .

PROBLEM NO.  1              TESTING DATA SCREENING AND NORMALITY TEST PROGRAM

SELECTION   9               VARIABLE NO.  9              DIAGNOSIS                   TRANSFORMATION NO.  4

DATA BEFORE TRANSFORMATION
  OBS. NO.

    1-   8   6.0000E 00  1.0000E 00  0.0         6.0000E 00  9.0000E 00  1.0000E 00  3.0000E 00  5.0000E 00
    9-  16   4.0000E 00  3.0000E 00  4.0000E 00  6.0000E 00  5.0000E 00  4.0000E 00  8.0000E 00  3.0000E 00
   17-  24   3.0000E 00  1.0000E 00  5.0000E 00  8.0000E 00  6.0000E 00  3.0000E 00  8.0000E 00  5.0000E 00
   25-  32   3.0000E 00  3.0000E 00  4.0000E 00  6.0000E 00  7.0000E 00  5.0000E 00  3.0000E 00  6.0000E 00
   33-  40   6.0000E 00  1.0000E 00  3.0000E 00  1.0000E 00  1.0000E 00  5.0000E 00  5.0000E 00  1.0000E 00
   41-  48   1.0000E 00  8.0000E 00  2.0000E 00  1.0000E 00  1.0000E 00  7.0000E 00  9.0000E 00  2.0000E 00
   49-  56   2.0000E 00  1.0000E 00  5.0000E 00  1.0000E 00  1.0000E 00  7.0000E 00  7.0000E 00  2.0000E 00
   57-  64   7.0000E 00  2.0000E 00  6.0000E 00  5.0000E 00  5.0000E 00  3.0000E 00  0.0         2.0000E 00
   65-  72   6.0000E 00  6.0000E 00  2.0000E 00  7.0000E 00  8.0000E 00  5.0000E 00  3.0000E 00  6.0000E 00
   73-  80   6.0000E 00  5.0000E 00  8.0000E 00  3.0000E 00  6.0000E 00  1.0000E 00  8.0000E 00  7.0000E 00
   81-  88   2.0000E 00  3.0000E 00  2.0000E 00  5.0000E 00  3.0000E 00  6.0000E 00  3.0000E 00  2.0000E 00
   89-  96   6.0000E 00  4.0000E 00  7.0000E 00  9.0000E 00  4.0000E 00  5.0000E 00  1.0000E 00  1.0000E 00
   97- 100   6.0000E 00  8.0000E 00  6.0000E 00  4.0000E 00

DATA AFTER  TRANSFORMATION
  OBS. NO.

    1-   8   2.4495E 00  1.0000E 00  0.0         2.4495E 00  3.0000E 00  1.0000E 00  1.7321E 00  2.2361E 00
    9-  16   2.0000E 00  1.7321E 00  2.0000E 00  2.4495E 00  2.2361E 00  2.0000E 00  2.8284E 00  1.7321E 00
   17-  24   1.7321E 00  1.0000E 00  2.2361E 00  2.8284E 00  2.4495E 00  1.7321E 00  2.8284E 00  2.2361E 00
   25-  32   1.7321E 00  1.7321E 00  2.0000E 00  2.4495E 00  2.6458E 00  2.2361E 00  1.7321E 00  2.4495E 00
   33-  40   2.4495E 00  1.0000E 00  1.7321E 00  1.0000E 00  1.0000E 00  2.2361E 00  2.2361E 00  1.0000E 00
   41-  48   1.0000E 00  2.8284E 00  1.4142E 00  1.0000E 00  1.0000E 00  2.6458E 00  3.0000E 00  1.4142E 00
   49-  56   1.4142E 00  1.0000E 00  2.2361E 00  1.0000E 00  1.0000E 00  2.6458E 00  2.6458E 00  1.4142E 00
   57-  64   2.6458E 00  1.4142E 00  2.4495E 00  2.2361E 00  2.2361E 00  1.7321E 00  0.0         1.4142E 00
   65-  72   2.4495E 00  2.4495E 00  1.4142E 00  2.6458E 00  2.8284E 00  2.2361E 00  1.7321E 00  2.4495E 00
   73-  80   2.4495E 00  2.2361E 00  2.8284E 00  1.7321E 00  2.4495E 00  1.0000E 00  2.8284E 00  2.6458E 00
   81-  88   1.4142E 00  1.7321E 00  1.4142E 00  2.2361E 00  1.7321E 00  2.4495E 00  1.7321E 00  1.4142E 00
   89-  96   2.4495E 00  2.0000E 00  2.6458E 00  3.0000E 00  2.0000E 00  2.2361E 00  1.0000E 00  1.0000E 00
   97- 100   2.4495E 00  2.8284E 00  2.4495E 00  2.0000E 00


 THE FOLLOWING CASES HAVE MISSING DATA. . .
      3
     63

 SAMPLE SIZE =    98
 MEAN =              1.9986
 STD. DEV. =        0.6138
 MEDIAN =              2.2361
 SKEW =           -1.1607

 EXTREMES GREATER THAN  1.000 OR LESS THAN  -1.000. . .
  OBS. NO.                 TRANSFORMED VALUE      STANDARDIZED VALUE


       2                      1.00000E 00               -1.6268
       5                      3.00000E 00                1.6314
       6                      1.00000E 00               -1.6268
      15                      2.82843E 00                1.3519
      18                      1.00000E 00               -1.6268
       .                           .                        .
       .                           .                        .
       .                           .                        .

      91                      2.64575E 00                1.0543
      92                      3.00000E 00                1.6314
      95                      1.00000E 00               -1.6268
      96                      1.00000E 00               -1.6268
      98                      2.82843E 00                1.3519
```

continued

```
PROBLEM NO.  1                TESTING DATA SCREENING AND NORMALITY TEST PRCGRAM
SELECTION   9                 VARIABLE NO.  9              DIAGNOSIS                  TRANSFCRMATICN NO.  4
FREQUENCY

   38 -

   36 -

   34 -

   32 -

   30 -

   28 -

   26 -                    X

   24 -

   22 -

   20 -
                                                                                                 X
   18 -
                                                                                       X
   16 -
                                               X
   14 -                                                                      X

   12 -

   10 -

    8 -
                                                                   X
    6 -

    4 -

    2 -

    0 -          X                   X                   X                                                 X
DECILES     I         I         I         I         I         I         I         I         I         I         I
         -INF     -1.282    -0.842    -0.524    -0.253        0    0.253     0.524     C.842     1.282     +INF
      N           0        26         0        15         0         7        14        17        19         0    TOTAL =   98
      +          0.0      16.2       0.0       5.2       0.0       0.0       4.2       7.2       9.2       0.0
      -         -9.8       0.0      -9.8       0.0      -9.8      -2.8       0.0       0.0       0.0      -9.8
   EXPECTED FREQUENCY =   9.8
   CHI-SQUARE =    85.2653        D. F. = 7
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $3.29.

Charge to user = computer costs + postage + network overhead
= $3.29 + postage + network overhead

CONTENTS—NORMSURV

pages
1– 2 Identification & Abstract
3– 6 User Instructions
7–14 I/O
15 Cost—Contents

DESCRIPTIVE TITLE	Cross-Cultural Comparison
CALLING NAME	POLYCOMP
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

POLYCOMP is designed for cross-cultural comparison using the data contained in Murdock's *Ethnographic Atlas*.[1] (If a copy of the "Ethnographic Atlas Codesheet"[2] is not immediately available, selected portions of it may be examined during the running of the program, or by running ETH–CODE, EIN No. 000 0119.) The user has the option of working with either one or two groups of cultures (or political units). Up to 20 cultures may be assigned to each group. A comparison is first made for each unit with each of the 92 characteristics from the *Ethnographic Atlas* compared for each culture against every other culture in that group. Characteristics with "insufficient information" are discounted.

The user then has the option of having the identity matrices printed out in one of two formats, after which he may weight the characteristics according to his own preference and obtain a new comparison table.

POLYCOMP may be used in conjunction with ETH–CODE, ETH–DGRE, ETH–RAND, ETH–INFO, ETHATLAS, CULTCOMP, and CULTPIK (EIN Nos. 000 0119–000 0125).

continued

REFERENCES

1. Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

3. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2, (April 1967).

4. "Ethnographic Atlas", by the editors of *Ethnology, I*, 1, (Jan., 1962)–*VII*, 3, (July, 1968), installments of information in each issue.

5. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

USER INSTRUCTIONS

In response to the NEW OR OLD query, type NEW ZILCH1 and press 'return'. When the computer responds READY, type SAVE and press 'return'. The computer will again respond READY, after which the user should type NEW ZILCH2 and SAVE it also when the computer is ready. Note: ALL INPUT SHOULD BE IN CAPITAL LETTERS, INCLUDING RESPONSES TO QUESTIONS.

When the above operations have been completed, call the program following the procedures outlined in Ref. 1. Type YES in response to the question "Have you saved ZILCH1 and ZILCH2?", and supply your name when asked (fictitious names are accepted). If you are using the program for the first time, type NO when asked if you are familiar with the format and options of the program. This will provide you with a two-page description of how to proceed through the program, which you should not have to repeat again.

When you are through running POLYCOMP, type UNSAVE ZILCH1 and press 'return'; when the computer is READY, type UNSAVE ZILCH2 and press 'return'. When the computer is READY again, you can sign off in the normal way. If you do not UNSAVE these files, you will incur a small storage charge for them.

REFERENCES

1. Morton, A.K., Concise Handbook to the Dartmouth Time-Sharing System, (Kiewit Comput. Center, Hanover, N.H., June 1970). Available from the EIN Office for the cost of reproduction and mailing.

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970). Revised. Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
RUN

P0LYC0MP  13 0CT 70  09:20

HAVE Y0U SAVED ZILCH1 AND ZILCH2? YES


Y0UR NAME? YURI MALIN0WSKI

-------------------------------------------------------------


                             -1-

P0LYC0MP                                      YURI MALIN0WSKI

ARE Y0U FAMILIAR, YURI, WITH THE F0RMAT 0F THIS PR0GRAM AND THE
   0PTI0NS WHICH Y0U MAY EXERCISE? N0

                        ===================
                        = P 0 L Y C 0 M P =
                        ===================

        THIS PR0GRAM IS DESIGNED F0R THE S0CIAL SCIENCE STUDENT
     ENGAGED IN CR0SS CULTURAL  RESEARCH.  THE  STUDENT  HAS  THE
     0PP0RTUNITY T0 C0MPARE TW0  GR0UPS  0F  CULTURES  0R  TRIBES
     (E.G., P0LITICAL UNITS) USING THE  DATA  FR0M  GE0RGE  PETER
     MURD0CK'S 'ETHN0GRAPHIC ATLAS.' THE STUDENT CAN THEN  WEIGHT
     THE 'ETHN0GRAPHIC ATLAS' CHARACTERISTICS  T0  REFLECT  THEIR
     RELATIVE IMP0RTANCE T0 THE HYP0THESIS HE IS STUDYING.

        AS AN EXAMPLE, SUPP0SE Y0U WERE INTERESTED IN C0MPARING
     THE P0LITICAL UNITY 0F  TW0  AFRICAN  C0UNTRIES,  NIGERIA  &
     BIAFRA, AND HAD ALREADY DECIDED THAT THERE WERE 13  CULTURES
     IN NIGERIA WHICH SH0ULD BE C0NSIDERED AND 8 IN BIAFRA.

        HAVING FIRST  SAVED  "ZILCH1"  AND  "ZILCH2,"  Y0U  RUN
     "P0LYC0MP***" AND SUPPLY THE AB0VE  INF0RMATI0N.  THERE  ARE
     1168 CULTURES REC0RDED UNDER THE 'ETHN0GRAPHIC ATLAS'  (THEY
     CAN BE LISTED BY REGI0N UNDER "CULTC0MP***"). H0WEVER,  S0ME
     CULTURES WHICH ARE 0F INTEREST T0 Y0U MAY N0T  BE  AVAILABLE
     0R ARE REC0RDED IN THE 'ATLAS' UNDER AN0THER  NAME.  ALL  0F
     THE CULTURES WHICH APPEAR IN  "ETH-INF0***,"  "ETH-DGRE***,"
     AND "ETH-RAND***" ARE N0T AVAILABLE T0 THIS PR0GRAM.

        THE FIRST TIME THE CULTURES WITHIN  A  DIVISI0N  (E.G.,
     NIGERIA, BIAFRA, 0R A RAND0M SAMPLE) ARE  C0MPARED,  ALL  0F
     THE 92 CHARACTERISTICS FR0M THE 'ETHN0GRAPHIC ATLAS' RECEIVE
     EQUAL WEIGHTING. THE USER RECEIVES A SUMMARY  0F  THE  T0TAL
     SIMILARITIES, DISSIMILARITIES,  AND  DISC0UNTED  C0MPARIS0NS
     ("INSUFFICIENT INF0" F0R EITHER CULTURE) F0R  EACH  DIVISI0N
     HE IS W0RKING WITH. THE USER  IS  THEN  PRESENTED  WITH  THE
     0PTI0N T0 SEE THE INDIVIDUAL  CULTURAL  C0MPARIS0NS  PRINTED
     0UT IN 0NE 0F TW0 F0RMATS F0R EACH DIVISI0N.  F0LL0WING  ARE
     EXAMPLES 0F THE TW0 F0RMATS:

          GRAPHIC:
                          Y                L              I
             YURAK        *                +              -
             LAPPS        +                *              +
             ICELANDERS   -                +              *

          NUMERIC:
                          Y                L              I
             YURAK                        -12            -1
             LAPPS        12                             21
             ICELANDERS   -1              21
```

continued

POLYCOMP (CONT.) YURI MALINOWSKI

CULTURES WITHIN THE DIVISION ARE LISTED ON THE LEFT OF THE IDENTITY MATRIX AND THEIR FIRST LETTERS ACROSS THE TOP. IN THE GRAPHIC FORMAT '+' SIGNIFIES SIMILARITIES > DISSIMILARITIES, '-' = DISSIMILARITIES > SIMILARITIES, '0' = SIMILARITIES = DISSIMILARITIES, '*' = THE IDENTITY ELEMENT.

THE 'ETHNOGRAPHIC ATLAS' IS THEN LISTED IF THE USER WISHES TO WEIGHT THE CHARACTERISTICS FOR A NEW SUMMARY. THE USER CAN ALSO LIST THE 'ATLAS' CODESHEET BY RUNNING "ETH-CODE***." THERE ARE 92 CHARACTERISTICS IN THE 'ATLAS.' FORTY-EIGHT OF THESE CHARACTERISTICS ARE CALLED 'MAJOR CHARACTERISTICS.' IN THE WEIGHTING PROCESS THE USER CAN ASSIGN WEIGHTS BY MAJOR CHARACTERISTIC (IN WHICH CASE ALL CHARACTERISTICS WHICH FALL UNDER THE MAJOR CHARACTERISTIC ARE ASSIGNED THE SAME WEIGHT), OR HE MAY ASSIGN WEIGHTS TO INDIVIDUAL COLUMNS, USING THE COLUMN CODE NUMBER WHICH APPEARS AT THE FAR LEFT OF THE CODESHEET PRINT-OUT. IN THE FOLLOWING EXAMPLE OF CODESHEET FORMAT, "FAMILY ORGANIZATION" IS A MAJOR CHARACTERISTIC WHICH IS NUMBERED 4 ON THE LIST OF MAJOR CHARACTERISTICS. "PREVAILING FORM OF DOMESTIC ORGANIZATION" IS A SUB-CHARACTERISTIC UNDER "FAMILY ORGANIZATION" WHICH WOULD BE REFERRED TO BY ITS COLUMN CODE OF 213.

```
                    <<FAMILY ORGANIZATION>>

213 PREVAILING FORM OF DOMESTIC ORGANIZATION
```

WEIGHTS ARE ASSIGNED IN FOUR GROUPS: 'VERY IMPORTANT,' 'IMPORTANT,' 'SLIGHTLY IMPORTANT,' AND UNIMPORTANT IF NOT SPECIFIED UNDER ONE OF THE ABOVE HEADINGS (IT IS THEN TREATED THE SAME AS "INSUFFICIENT INFO" AND DISCOUNTED). CAREFUL ATTENTION SHOULD BE GIVEN TO A STUDY OF THE CODESHEET AND THE ASSIGNING OF WEIGHTS.

FOR INFORMATION ON PROGRAMS RELATED TO THIS AND PROGRAMS GIVING INFORMATION, RANDOM SAMPLES, AND OTHER COMPARISONS, LIST "ANTHRO***."

```
ARE YOU WORKING WITH 1 OR 2 CULTURAL OR TRIBAL GROUPINGS? 1
WHAT IS THE NAME OF THE GROUP YOU ARE WORKING WITH? SLAVS
HOW MANY CULTURES FALL UNDER THE SLAVS DIVISION? 4
          TYPE IN THE NAMES OF THE 4  CULTURES . . .

                          --SLAVS--

? ARMENIANS,BYELORUSSI,CZECHS,SERBS




                                                     --GROSS CALCULATION--
                                                 SIMILAR-DISSIMILAR-DISCOUNTED

SLAVS                                              319       131         102

WOULD YOU LIKE TO SEE A PRINT-OUT OF THE SIMILARITIES
   & DISSIMILARITIES FOR SLAVS? YES

WOULD YOU RATHER SEE THE DIFFERENCE BETWEEN THE SIMILARITIES
   & DISSIMILARITIES AS A NUMBER OR IN A GRAPHIC REPRESENTATION.
   TYPE EITHER A 1 OR A 2 TO INDICATE YOUR PREFERENCE? 1
```

continued

7

```
POLYCOMP (CONT.)                                  YURI MALINOWSKI


         FORMAT:                                 'CULTURE'
'CULTURE'          SIMILARITIES MINUS DISSIMILARITIES

                              --SLAVS--

                       A            B            C            S
ARMENIANS                           31           33           21
BYELORUSSI             31                        48           20
CZECHS                 33           48                        35
SERBS                  21           20           35


DO YOU WANT TO WEIGHT THE CHARACTERISTICS FROM THE
    'ETHNOGRAPHIC ATLAS' CODESHEET AND OBTAIN A NEW SET
    OF COMPARISONS? YES
DO YOU WANT TO SEE THE 'ETHNOGRAPHIC ATLAS' CODESHEET? NO
IN WEIGHTING THE 'ETHNOGRAPHIC ATLAS' CHARACTERISTICS, YOU
     HAVE THE OPPORTUNITY TO SPECIFY HOW IMPORTANT CERTAIN
     CHARACTERISTICS ARE FOR YOUR OWN STUDY. YOU MAY ASSIGN
     WEIGHTS BY GROUP AS TO THOSE CHARACTERISTICS WHICH ARE
     'VERY IMPORTANT,' 'IMPORTANT,' OR 'SLIGHTLY IMPORTANT.' IF
     IF YOU DO NOT ASSIGN A CHARACTERISTIC TO ONE OF THESE
     THREE GROUPS IT WILL BE ASSUMED THAT IT IS UNIMPORTANT,
     AND THEREFORE DISCOUNTED.

AFTER EACH WEIGHTING CATEGORY YOU WILL BE ASKED TO SUPPLY THE
     'CHARACTERISTIC NUMBERS' FIRST--THESE NUMBERS APPEARED
     TO THE LEFT OF THE LIST OF 48 MAJOR CHARACTERISTICS.
     RIGHT AFTER THAT YOU WILL BE ASKED FOR 'COLUMN NUMBERS.'
     IF YOU LISTED ANY SEGMENTS OF THE CODESHEET, YOU WILL NOTE
     THAT SOME OF THE MAJOR CHARACTERISTICS ARE SUBDIVIDED.
     IF YOU WANT TO ASSIGN THESE SUBDIVISIONS SPECIAL WEIGHTS
     YOU MAY TYPE IN THEIR COLUMN NUMBERS (MAKE SURE YOU HAVE
     THE COLUMN NUMBERS AND NOT THE CODES!).

IF YOU DO NOT HAVE ANY CHARACTERISTICS YOU WISH TO ASSIGN A
     CERTAIN WEIGHT, OR YOU DO NOT HAVE ANY COLUMN NUMBERS,
     SIMPLY PRESS 'RETURN.'


     MOST IMPORTANT CHARACTERISTICS . . .
        CHARACTERISTIC NUMBERS? 12,32,33
        COLUMN NUMBERS?


     IMPORTANT CHARACTERISTICS . . .
        CHARACTERISTIC NUMBERS? 46,47
        COLUMN NUMBERS? 27


     SLIGHTLY IMPORTANT CHARACTERISTICS . . .
        CHARACTERISTIC NUMBERS? 27
        COLUMN NUMBERS?



                                            --GROSS CALCULATION--
                                         SIMILAR-DISSIMILAR-DISCOUNTED

SLAVS                                       77        36        507

WOULD YOU LIKE TO SEE A PRINT-OUT OF THE SIMILARITIES
   & DISSIMILARITIES FOR SLAVS? YES

WOULD YOU RATHER SEE THE DIFFERENCE BETWEEN THE SIMILARITIES
   & DISSIMILARITIES AS A NUMBER OR IN A GRAPHIC REPRESENTATION.
   TYPE EITHER A 1 OR A 2 TO INDICATE YOUR PREFERENCE? 1
```

continued

```
POLYCOMP (CONT.)                              YURI MALINOWSKI

        FORMAT:                          'CULTURE'
'CULTURE'  MOST IMPORTANT -->  SIMILARITIES MINUS DISSIMILARITIES
           IMPORTANT -->            ''          ''       ''
           SLIGHTLY IMPORTANT -->   ''          ''       ''

                              --SLAVS--

                          A             B             C             S
ARMENIANS                               3             1            -1
                                        1             1             1
                                        1            -1            -1
BYELORUSSI                3                           3             1
                          1                           1             1
                          1                          -1            -1
CZECHS                    1             3                           3
                          1             1                           1
                         -1            -1                           2
SERBS                    -1             1             3
                          1             1             1
                         -1            -1             2

DO YOU WANT TO GO THROUGH THE WEIGHTING AGAIN? NO


- ----------------------------------------------------------------------

TIME: 5.600 SEC.
```

FILMED FROM BEST AVAILABLE COPY

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 5.6 seconds, at a rate of $0.11/sec. Total terminal time was about 25 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead

= $0.62 + $1.45 + communications + network overhead

= $2.07 + communications + network overhead

CONTENTS—POLYCOMP

pages	
1– 2	Identification & Abstract
3	User Instructions
5– 8	I/O
9	Cost—Contents

DESCRIPTIVE TITLE	Major characteristics from Murdock's *Ethnographic Atlas*
CALLING NAME	ETH—CODE
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Codesheet revised from that of Herbert Barry III, Department of Anthropology, University of Pittsburgh
LANGUAGE	Dartmouth BASIC
COMPUTER	GE—635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETH—CODE is a driver program for the sub-program CODESHEET, which contains the 48 major characteristics from the *Ethnographic Atlas*[1-5]. After these have been listed, the user has the option to see the individual codings and subdivisions for any of these major characteristics. The output takes the form of a list with matching column and code numbers as they appear in "Ethnographic Atlas Codesheet"[2]. This program is, in effect, a partial replacement for the complete codesheet, if the user does not have one available. Requests to have individual codings output should be limited to only a few of the major characteristics, as the complete codesheet produces many pages of output and requires considerable time.

REFERENCES

1. Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

continued

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

3. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

4. "Ethnographic Atlas", by the editors of *Ethnology*, *I*, 1(Jan., 1962)–*VII*, 3,(July, 1968), installments of information in each issue.

5. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology*, *VI*, 2,(April, 1967).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook"[1]. After the 48 major characteristics have been listed, you will be asked if you wish to see individual codings; if so, supply the numbers of interest to you when asked.

REFERENCES

1. Morton, A.K., *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., June, 1970). Available from the EIN Office at the cost of reproduction and mailing.

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
MAJOR CHARACTERISTICS FROM THE 'ETHNOGRAPHIC ATLAS'


              1  LOCATION
              2  SUBSISTENCE ECONOMY
              3  MODE OF MARRIAGE
              4  FAMILY ORGANIZATION
              5  MARITAL RESIDENCE
              6  COMMUNITY ORGANIZATION
              7  PAT MAT KIN GROUPS & EXOGAMY
              8  COGNATIC KIN GROUPS
              9  COUSIN MARRIAGE
             10  FIRST COUSIN TERMINOLOGY
             11  TYPE AND INTENSITY OF AGRICULTURE
             12  SETTLEMENT PATTERN
             13  MEAN SIZE OF LOCAL COMMUNITIES
             14  JURISDICTIONAL HIERARCHY
             15  HIGH GODS
             16  TYPES OF GAMES
             17  POST-PARTUM SEX TABOOS
             18  MALE GENITAL MUTILATIONS
             19  SEGREGATION OF ADOLESCENT BOYS
             20  TYPE OF ANIMAL HUSBANDRY
             21  SUMMARY CODES
             22  METAL WORKING
             23  WEAVING
             24  LEATHER WORKING
             25  POTTERY
             26  BOAT BUILDING
             27  HOUSE CONSTRUCTION
             28  GATHERING
             29  HUNTING
             30  FISHING
             31  ANIMAL HUSBANDRY
             32  AGRICULTURE
             33  LINGUISTIC AFFILIATION
             34  CLASS STRATIFICATION
             35  CASTE STRATIFICATION
             36  SLAVERY
             37  SUCCESSION OF LOCAL HEADSMAN
             38  INHERITANCE OF REAL PROPERTY
             39  INHERITANCE OF MOVABLE PROPERTY
             40  NORMS FOR GIRL'S PREMARITAL SEX
             41  GROUND PLAN OF DWELLING
             42  FLOOR PLAN OF DWELLING
             43  WALL MATERIAL
             44  SHAPE OF ROOF
             45  ROOFING MATERIAL
             46  POLITICAL INTEGRATION
             47  POLITICAL SUCCESSION OF LOCAL COMMUNITY
             48  ENVIRONMENT



WOULD YOU LIKE TO SEE THE INDIVIDUAL CODINGS AND SUBDIVISIONS
       FOR ANY OF THESE 48 MAJOR CHARACTERISTICS? YES

TYPE THE NUMBERS OF THE MAJOR CHARACTERISTICS WHICH YOU
       WOULD LIKE TO SEE IN FULL? 14,15,33


COLUMN    CODE


                    <<JURISDICTIONAL HIERARCHY>>

 39  LOCAL COMMUNITY (NUCLEAR FAMILY, EXTENDED
                     FAMILY, CLAN-BARRIO, VILLAGE)
```

continued

```
SELECTIONS FROM THE ETHNOGRAPHIC ATLAS (CONT.)

COLUMN   CODE


  40  BEYOND LOCAL COMMUNITY (PARISH, DISTRICT,
                    PROVINCE, NATION-STATE)

            0     INSUFFICIENT INFORMATION
            1     NO LEVELS
            2     ONE LEVEL
            3     TWO LEVELS
            4     THREE LEVELS
            5     FOUR LEVELS


  41                         <<HIGH GODS>>

            0     INSUFFICIENT INFORMATION
            1     ABSENT OR NOT REPORTED
            2     PRESENT BUT NOT ACTIVE IN HUMAN AFFAIRS
            3     PRESENT AND ACTIVE IN HUMAN AFFAIRS-NOT SUPPOR-
                           TIVE OF HUMAN MORALITY
            4     PRESENT ACTIVE & SPECIFICALLY SUPPORTIVE OF
                           HUMAN MORALITY


  272                   <<LINGUISTIC AFFILIATION>>

      272         PHYLUM

      274         FAMILY
               1 NIGER-CONGO                 2     MALAYO-POLYNESIAN
               3 THAI-KADAI                  4     ANNAM-MUONG
               5 MON-KHMER                   6     INDO-EUROPEAN
               7 AFRO-ASIATIC                8     ALGONKIAN
               9 RITWAN                      10    ALTAIC
               11 URALIC                     12    ATHAPASKAN
               13 CARIBAN                    14    PEBAN
               15 CHARI-NILE                 16    TIBETO-BURMAN
               17 SINITIC                    18    PENUTIAN
               19 SAHAPTIN                   20    OREGON PENUTIAN
               21 YAKONAN                    22    HOKAN
               23 TANOAN                     24    NAHUATLAN
               25 PIMAN                      26    SHOSHONEAN
               27 TARACAHITIAN               28    TIPI-GUARANI
               29 WITOTAN                    30    ZAPAROAN
               31 CHEMAKUAN                  32    SALISHAN
               33 WAKASHAN                   34    ABASGO-KERKETIAN
               35 CHENCHENO-LESGHIAN         36    GEORGIAN
               37 BORORAN                    38    CAINGANG
               39 GE                         40    CHIBCHAN
               41 MISUMALPAN                 42    CHIQUITOAN
               43 GUAYCURAN                  44    MATACO-MATE GUAYO
               45 AUSTRALIAN                 46   -PAPUAN
               47 MIXTECAN                   48    ZAPOTECAN
               49 ARAWAKAN                   50    ARAUCANIAN
               51 BETOYAN                    52    CADDOAN
               53 CAHUAPANAN                 54    DRAVIDIAN
               55 ESKIMAUAN                  56    IROQUOIAN
               57 JAPANO-RYUKUAN             58    KANURIC
               59 KECHUMARAN                 60    KHOISAN
               61 KATUKINAN                  62    KOMAN
               63 KORDOFANIAN                64    KERESAN
               65 LUORAWETLAN                66    MAYAN
               67 MASCOIAN                   68    MIAO-YAO
               69 MIZOCUAVEAN                70    NATCHEZ-MUSKOGEAN
               71 OTO-MAGUEAN                72    PANOAN
               73 PUINAVEAN                  74    SIOUAN
               75 TEHUELCHEAN                76    TUNICAN
               77 YUNGA-PURUHAN              78    YUKIAN
               79 ZAMUCOAN
               80 ISOLATED FAMILY OR ISOLATED FAMILY PRESUMED
                 TO BE PART OF A PHYLUM

      276         SUBFAMILY


TIME: 1.982 SEC.
```

COST ESTIMATE

For the job included as the sample output, the total central processor unit (CPU) time was 1.98 seconds, at a rate of \$0.11/sec. Total terminal time was about 10 minutes, at a rate of \$3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead
= \$0.22 + \$0.58 + communications + network overhead
= \$0.80 + communications + network overhead

CONTENTS--ETH-CODE

pages

1- 2 Identification & Abstract
3 User Instructions
5- 6 I/O
7 Cost--Contents

EDUCOM EDUCATIONAL INFORMATION NETWORK

000 0120

DESCRIPTIVE TITLE	Random Samples of Cultures
CALLING NAME	ETH—DGRE
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data collected by Daniel Gordon, Department of Sociology, University of Oregon, from Murdock's *Ethnographic Atlas*
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

One of the chief problems in making cultural comparisons and gathering a random sample of cultures is making sure they are not so close together as to share many characteristics due to diffusion and common environmental pressures. In order to avoid this, cultural anthropologists employ a 3—degree rule, to insure a geographical separation. Respecting this rule, ETH—DGRE randomly selects one culture (if any cultures have been recorded[1] to exist in the sample) in each of the 72 cells of 30—degree latitude and longitude on mercator projection. Each selected culture is then compared with the cultures of the neighboring cells to check for the 3—degree rule. This includes checking cultures of the far west with cultures of the far east, and those of the far north with those of the far south, when such cultures exist.

Output is in the form of a simple list or a two-page map with 30—degree cells indicated.

continued

The user cannot be guaranteed to get a sample of given size by using ETH—DGRE, whereas ETH—RAND (EIN No. 000 0121) will supply a sample of 50 cultures (and consume considerably more time in doing so).

Other programs which might be useful in conjunction with this one include ETH—INFO, ETH—CODE, ETHATLAS, CULTCOMP, CULTPIK, and POLYCOMP, (EIN No. 000 0118 - 000 0119, 000 0122 - 000 0125).

REFERENCES

1. "Condensed Geographic Codesheet of the Ethnographic Atlas: 861 Cultures", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970).

2. McNett, C.W., and Kirk, R.E., "Drawing Random Samples in Cross-Cultural Studies; A Suggested Method", Amer. Anthrop., *70*, 1, (Feb. 1968).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook".[1] Type 'list' or 'map' in response to the question of preferred form of output.

REFERENCES

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., June 1970). Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
ETH-DGRE  16 SEP 70  11:14

DO YOU WANT A SIMPLE LIST OF THE CULTURES, OR WOULD YOU
LIKE A TWO-PAGE MAP WITH LONGITUDINAL AND LATITUDINAL
COORDINATES (TYPE: 'LIST' OR 'MAP')? MAP

-----------------------------------------------------------------



RANDOM SAMPLE OF 861          - 1 -
  CULTURES SEPARATED
  BY THE 3 DEGREE RULE


NORTHWEST                                                        GREENWICH

     180         150         120          90          60          30           0
  90 +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +

     .           .           .           .           .           .           .
       TAKEUMIUT   CHUGACH     COPPER ESK  POLAR ESKI  ANGMAGSALI  ICELANDERS
          (5)         (6)         (5)         (3)         (1)         (1)
     .           .           .           .           .           .           .

  60 +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +

     .           .           .           .           .           .           .
       ALEUT       SHUSWAP     PLAINS CRE  WINNEBAGO     ******    IRISH
          (1)         (63)        (99)        (22)        (0)         (5)
     .           .           .           .           .           .           .

  30 +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +

     .           .           .           .           .           .           .
       HAWAIIANS     ******    TOTONAC     CHIBCHA     WAI WAI     BOBO
E         (1)         (0)         (18)        (34)        (6)         (50)
Q    .           .           .           .           .           .           .
U
A  0 +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +
T
O    .           .           .           .           .           .           .
R      MANIHIKIAN  MANGAREVAN    ******    CHIRIGUANO  BOTOCUDO      ******
          (10)        (3)         (0)         (16)        (25)        (0)
     .           .           .           .           .           .           .

  30 +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +

     .           .           .           .           .           .           .
         ******      ******      ******    ONA         BRAZILIANS  TRISTAN
          (0)         (0)         (0)         (5)         (1)         (1)
     .           .           .           .           .           .           .

  60 +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +

     .           .           .           .           .           .           .
         ******      ******      ******      ******      ******      ******
          (0)         (0)         (0)         (0)         (0)         (0)
     .           .           .           .           .           .           .

  90 +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +
     180         150         120          90          60          30           0

                                                                 GREENWICH
SOUTHWEST
```

continued

```
RANDOM SAMPLE OF 861           - 2 -
  CULTURES SEPARATED
  BY THE 3 DEGREE RULE


GREENWICH                                                      NORTHEAST

0           30          60          90          120         150         180
+   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   + 90

.           .           .           .           .           .           .
 LAPPS         ******    YURAK       KET         YAKUT       KORYAK
    (1)         (0)         (2)         (1)         (2)         (2)
.           .           .           .           .           .           .

+   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   + 60

.           .           .           .           .           .           .
 ALGERIANS  TURKS       PAHARI      NATCHEZ     JAPANESE      ******
    (15)        (22)        (10)        (7)         (5)         (0)
.           .           .           .           .           .           .

+   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   + 30

.           .           .           .           .           .           .
 KANEMBU    DIDINGA     SINDHI      GARO        SUGBUHANON PONAPEANS
    (88)        (66)        (20)        (41)        (21)        (7)        E
.           .           .           .           .           .           .  Q
                                                                           U
+   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   + 0  A
                                                                           T
.           .           .           .           .           .           .  O
 KUNG       LUO            ******   MACASSARES TORADJA     MOTA            R
    (34)        (53)        (0)         (7)         (46)        (26)
.           .           .           .           .           .           .

+   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   + 30

.           .           .           .           .           .           .
   ******   PONDO          ******      ******   WONGAIBON  MAORI
    (0)         (1)         (0)         (0)         (2)         (1)
.           .           .           .           .           .           .

+   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   + 60

.           .           .           .           .           .           .
   ******      ******      ******      ******      ******      ******
    (0)         (0)         (0)         (0)         (0)         (0)
.           .           .           .           .           .           .

+   -   -   +   -   -   +   -   -   +   -   -   +   -   -   +   -   -   + 90
0           30          60          90          120         150         180

GREENWICH                                                      SOUTHEAST



-----------------------------------------------------------------------

TIME: 2.150 SEC.
```

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 2.15 seconds, at a rate of $0.11/sec. Total terminal time was about 9 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead
= $0.24 + $0.53 + communications + network overhead
= $0.77 + communications + network overhead

CONTENTS—ETH-DGRE

pages

1– 2 Identification & Abstract
3 User Instructions
5– 6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Random Sample of 50 Cultures
CALLING NAME	ETH-RAND
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data collected by Daniel Gordon, Department of Sociology, University of Oregon, from Murdock's *Ethnographic Atlas*
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETH-RAND provides an alternative to ETH-DGRE (EIN No. 000 0120) in randomly selecting a sample of 50 cultures separated by the 3-degree rule. The program initially selects a random point of specific latitude and longitude. With this point as an origin the program then lays out a mercator projection and selects 50 points by randomly choosing latitude and longitude points which are then converted to points on the normal mercator projection map. ETH-RAND then examines the 30-degree cell in which this point lies and the eight surrounding cells to determine the culture closest to this point. This culture must pass the 3-degree rule or the program persists until it has found the closest culture which does pass the test. The program then prints out the coordinates of the random point, followed by the selected culture and its coordinates.

continued

The program CULTCOMP (EIN No. 000 0124) and ETH-INFO (EIN No. 000 0122) may be employed to gather additional information on these cultures.

REFERENCES

"Condensed Geographic Code Sheet of the Ethnographic Atlas: 861 Cultures", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970).

McNett, C.W., and Kirk, R.E., "Drawing Random Samples in Cross-Cultural Studies; A Suggested Method", Amer. Anthrop., *70*, 1, (Feb., 1968).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook".[1] There is nothing further to do in this program.

References

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., June 1970). Available from the EIN Office at the cost of reproduction and mailing.

000 0121

SAMPLE INPUT AND OUTPUT

```
RUN

ETH-RAND  16 SEP 70  11:25

----------------------------------------------------------------------


RANDOM SAMPLE OF 50 CULTURES  - 1 -


  --RANDOM GRID--                          --NEAREST CULTURE--
LATITUDE    LONGITUDE               CULTURE     LATITUDE      LONGITUDE
======================================================================
 11 SOUTH    116 WEST              MANGAREVAN   20 SOUTH     134 WEST
 21 SOUTH    171 WEST              NIUEANS      19 SOUTH     169 WEST
 34 SOUTH    131 EAST              DIERI        28 SOUTH     138 EAST
 9  SOUTH    16  WEST              SHERBRO      7  NORTH     13  WEST
 14 SOUTH    176 EAST              ROTUMANS     13 SOUTH     177 EAST
 64 NORTH    141 WEST              NABESNA      63 NORTH     141 WEST
 4  SOUTH    121 EAST              TORADJA      2  SOUTH     121 EAST
 19 NORTH    81  EAST              MARIA GOND   19 NORTH     81  EAST
 24 SOUTH    96  EAST              MENTAWEIAN   3  SOUTH     100 EAST
 9  SOUTH    56  EAST              MERINA       19 SOUTH     46  EAST
 69 NORTH    66  WEST              BAFFINLAND   65 NORTH     65  WEST
 59 NORTH    136 EAST              GILYAK       53 NORTH     142 EAST
 54 SOUTH    31  WEST              BRAZILIANS   43 SOUTH     47  WEST
 19 NORTH    176 WEST              HAWAIIANS    20 NORTH     156 WEST
 34 NORTH    171 EAST              BIKINIANS    12 NORTH     165 EAST
 9  SOUTH    161 WEST              MANIHIKIAN   10 SOUTH     160 WEST
 54 SOUTH    41  EAST              PONDO        31 SOUTH     30  EAST
 11 SOUTH    81  WEST              CAMPA        8  SOUTH     75  WEST
 21 SOUTH    121 WEST              RAROIANS     16 SOUTH     142 WEST
 29 NORTH    151 EAST              CAROLINIAN   15 NORTH     146 EAST
 16 SOUTH    176 WEST              EASTER       27 SOUTH     190 WEST
 19 NORTH    51  EAST              MUTAIR       28 NORTH     47  EAST
 89 NORTH    151 EAST              YUKAGHIR     70 NORTH     145 EAST
 29 NORTH    86  EAST              SHERPA       28 NORTH     87  EAST
 44 NORTH    151 EAST              AINU         44 NORTH     144 EAST
 29 NORTH    161 EAST              TRUKESE      7  NORTH     152 EAST
 74 NORTH    41  WEST              ANGMAGSALI   66 NORTH     37  WEST
 19 SOUTH    151 WEST              TAHITIANS    18 SOUTH     150 WEST
 39 SOUTH    46  EAST              THONGA       24 SOUTH     32  EAST
 21 SOUTH    16  WEST              TRISTAN      37 SOUTH     12  WEST
 49 NORTH    6   WEST              IRISH        53 NORTH     9   WEST
 21 SOUTH    16  EAST              HERERO       21 SOUTH     16  EAST
 21 SOUTH    161 EAST              MAILU        10 SOUTH     149 EAST
 29 SOUTH    146 WEST              MANGAINAS    22 SOUTH     158 WEST
 74 NORTH    136 WEST              TLINGIT      58 NORTH     134 WEST
 29 SOUTH    171 EAST              MAORI        35 SOUTH     175 EAST
 14 NORTH    166 EAST              KUSAIANS     5  NORTH     163 EAST
 74 NORTH    6   WEST              ICELANDERS   64 NORTH     20  WEST
 6  SOUTH    76  EAST              SINHALESE    7  NORTH     80  EAST
 11 SOUTH    161 EAST              KAPINGAMAR   1  NORTH     155 EAST
 21 SOUTH    161 EAST              WONGAIBON    32 SOUTH     146 EAST
 59 SOUTH    111 WEST              ALACALUF     52 SOUTH     74  WEST
 14 NORTH    126 EAST              SUGBUHANON   10 NORTH     124 EAST
 29 NORTH    81  WEST              TIMUCUA      27 NORTH     82  WEST
 34 NORTH    1   WEST              RIFFIANS     35 NORTH     4   WEST
 39 SOUTH    141 WEST              MARQUESANS   9  SOUTH     140 WEST
 79 NORTH    91  WEST              IGLULIK      70 NORTH     82  WEST
 19 SOUTH    1   EAST              FANG         1  NORTH     11  EAST
 49 SOUTH    51  WEST              ONA          54 SOUTH     69  WEST
 64 NORTH    11  WEST              DUTCH        53 NORTH     7   EAST



----------------------------------------------------------------------

TIME: 33.994 SEC.
```

000 0121

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 34 seconds, at a rate of $0.11/sec. Total terminal time was about 8 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead
= $3.74 + $.46 + communications + network overhead
= $4.20 + communications + network overhead

CONTENTS—ETH-RAND

pages	
1– 2	Identification & Abstract
3	User Instructions
5	I/O
7	Cost—Contents

DESCRIPTIVE TITLE	Demographic Information from Murdock's *Ethnographic Atlas*
CALLING NAME	ETH-INFO
INSTALLATION NAME	Kiewit Computation Center Dartmouth College
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data collected by Daniel Gordon, Department of Sociology, University of Oregon, from Murdock's *Ethnographic Atlas*
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETH-INFO provides information on any of 861 cultures contained in the random access file DEGRE. Output consists of a table which shows the culture(s) of interest to the user, regional identification, latitude, longitude, size of population, year data were collected, and the year the population estimate was made. Definitions of regional identifications are output on request.

Other programs which might be useful in conjunction with ETH-INFO include ETH-CODE, ETH-DGRE, ETH-RAND, ETHATLAS, CULTCOMP, CULTPIK, and POLYCOMP, (EIN Nos. 000 0118 - 000 0125).

REFERENCES

Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

continued

000 0122

Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

"Ethnographic Atlas", by the editors of *Ethnology, I*, 1,(Jan., 1962)–*VII*, 3, (July, 1968), installments of information in each issue.

Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2, (April, 1967).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook".[1] Input the names of the cultures of interest to you when requested, following the instructions as they appear in the program.

The "Condensed Geographic Codesheet"[2] should be consulted for a list of the 861 cultures for which information is stored.

REFERENCES

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., June, 1970). Available from the EIN Office at the cost of reproduction and mailing.

2. "Condensed Geographic Codesheet of the Ethnographic Atlas: 861 Cultures" (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970).

SAMPLE INPUT AND OUTPUT

```
ETH-INFO  16 SEP 70  11:34
------------------------------------------------------------------


                              -1-

ETHNOGRAPHIC ATLAS INFORMATION


THIS PROGRAM PROVIDES INFORMATION FROM MURDOCK'S 'ETHNOGRAPHIC
  ATLAS' ON ANY ONE OF 861 CULTURES. THIS INCLUDES IDENTIFICATION,
  LATITUDE, LONGITUDE, TIME LEVEL TO WHICH DATA PERTAINS, POPULA-
  TION (TO NEAREST HUNDRED), AND THE YEAR OF POPULATION ESTIMATE
  (TO THE NEAREST DECADE).


FOR WHAT CULTURE OR CULTURES WOULD YOU LIKE INFORMATION . . .
   (IF YOU NEED MORE THAN ONE LINE TYPE ', AND' AND PRESS 'RETURN'.
? BLACKFOOT,CHEROKEE,CHEYENNE,CHOCTAW,CREEK,CROW



DO YOU WANT TO SEE THE DEFINITIONS FOR THE REGIONAL AREAS? YES

                        --IDENTIFICATION--

AFRICA:    EXCLUSIVE OF MADAGASCAR AND THE NORTHERN AND NORTH-
    EASTERN PORTIONS OF THE CONTINENT.
CIRCUM-MEDITERRANEAN:   INCLUDING EUROPE, TURKEY AND THE
    CAUCASUS, THE SEMITIC NEAR EAST, AND NORTHERN AND NORTH-
    EASTERN AFRICA.
EAST EURASIA:   EXCLUDING FORMOSA, THE PHILIPPINES, INDONESIA,
    AND THE AREA ASSIGNED TO THE CIRCUM-MEDITERRANEAN BUT
    INCLUDING MADAGASCAR AND OTHER ISLANDS IN THE INDIAN OCEAN.
INSULAR PACIFIC:   EMBRACING ALL OF OCEANIA AS WELL AS AREAS
    LIKE AUSTRALIA, INDONESIA, FORMOSA, AND THE PHILIPPINES
    THAT ARE NOT ALWAYS INCLUDED THEREWITH.
NORTH AMERICA:   INCLUDING THE INDIGENOUS SOCIETIES OF THIS
    CONTINENT AS FAR SOUTH AS THE ISTHMUS OF TEHUANTEPEC.
SOUTH AMERICA:   INCLUDING THE ANTILLES, YUCATAN, AND CENTRAL
    AMERICA AS WELL AS THE CONTINENT ITSELF.


==================================================================
CULTURE           *             LATITUDE * TIME OF DATA *   DATE OF EST.
     *       REGIONAL ID           * LONGITUDE *   POPULATION     *
 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
BLACKFOOT    NORTH AMERICA              051N 112W  1850    1600       1810
CHEROKEE     NORTH AMERICA              036N 083W  1750    20000      1730
CHEYENNE     NORTH AMERICA              039N 104W  1860    4700       1870
CHOCTAW      NORTH AMERICA              033N 088W  1760    15000      1760
CREEK        NORTH AMERICA              037N 085W  1820    1200       1930
CROW         NORTH AMERICA              045N 108W  1870    1800       1940
==================================================================


DO YOU WANT INFORMATION ON ANY OTHER CULTURES? YES
```

continued

```
ETHNOGRAPHIC INFORMATION (CONT.)


FOR WHAT CULTURE OR CULTURES WOULD YOU LIKE INFORMATION . . .
   (IF YOU NEED MORE THAN ONE LINE TYPE ', AND' AND PRESS 'RETURN'.
? ARMENIANS,BYELORUSSI,CZECHS,HUNGARIANS,SERBS



==================================================================
CULTURE           *                LATITUDE * TIME OF DATA *   DATE OF EST.
     *        REGIONAL ID          * LONGITUDE *   POPULATION   *
 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
ARMENIANS    CIRCUM-MEDITERRANEAN  040N 045E  1900    3400000    1950
BYELORUSSI   CIRCUM-MEDITERRANEAN  055N 028E  1910
CZECHS       CIRCUM-MEDITERRANEAN  050N 016E  1940    8800000    1950
HUNGARIANS   CIRCUM-MEDITERRANEAN  047N 020E  1940    9400000    1950
SERBS        CIRCUM-MEDITERRANEAN  044N 020E  1950
==================================================================

DO YOU WANT INFORMATION ON ANY OTHER CULTURES? NO




------------------------------------------------------------------
TIME: 9.771 SEC.
```

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 9.77 seconds, at a rate of \$0.11/sec. Total terminal time was about 7 minutes, at a rate of \$3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead
= \$1.08 + \$0.42 + communications + network overhead
= \$1.50 + communications + network overhead

CONTENTS—ETH-INFO

pages
1– 2 Identification & Abstract
3 User Instructions
5– 6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Comparison of Characteristics from "Ethnographic Atlas"
CALLING NAME	ETHATLAS
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	James Fernandez, Department of Anthropology, Dartmouth College Modified by Mark Hebenstreit, Class of 1970 Data coded by Herbert Barry III, Department of Anthropology, University of Pittsburgh
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

ETHATLAS scans data from the "Ethnographic Atlas" (92 characteristics of 1168 societies). It will compare any two characteristics on a presence or absence basis and compute degrees and significance of association. Any five conditions can be held constant. Thus, for example, the association may be computed for only those societies with intense agriculture or with patrilineal descent.

The program may be run only in consultation with the "Ethnographic Atlas Codesheet"[1] obtainable from the Department of Anthropology. If a codesheet is not immediately available, selected portions of it may be examined by running ETH–CODE, (EIN No. 000 0119). The codesheet gives the column numbers of the various characteristics and the coding relevant to decisions as to their presence or absence.

continued

000 0123

REFERENCES

1. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.
2. Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).
3. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).
4. "Ethnographic Atlas", by the editors of *Ethnology*, *I*, 1, (Jan., 1962)-*VII*, 3, (July, 1968), installments of information in each issue.
5. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology*, *VI*, 2, (April, 1967).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook".[1] Supply information on characteristics and codes as requested.

REFERENCES

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, June, 1970). Available from the EIN Office for the cost of reproduction and mailing.

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the Depart. of Anthrop. or from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
DØ YØU WANT PRØGRAM DESCRIPTIØN? YES ØR NØ? YES

          THIS PRØGRAM SCANS DATA FRØM THE ETHNØ-
          GRAPHIC ATLAS (92 CHARACTERISTICS ØF
          1168 SØCIETIES).  IT WILL CØMPARE ANY
          TWØ CHARACTERISTICS ØN A PRESENCE ØR
          ABSENCE BASIS AND CØMPUTE DEGREES AND
          SIGNIFICANCE ØF ASSØCIATIØN.  ANY FIVE
          CØNDITIØNS CAN BE HELD CØNSTANT. THUS,
          FØR EXAMPLE, THE ASSØCIATIØN MAY BE
          CØMPUTED FØR ØNLY THØSE SØCIETIES WITH
          INTENSE AGRICULTURE ØR WITH PATRILINEAL
          DESCENT.

          THE PRØGRAM CAN BE RUN ØNLY IN CØNSULTA-
          TIØN WITH A CØDESHEET ØBTAINABLE AT THE
          DEPARTMENT ØF ANTHRØPØLØGY.  IF A CØDE-
          SHEET IS NØT IMMEDIATELY AVAILABLE,
          SELECTED PØRTIØNS ØF IT MAY BE EXAMINED
          BY RUNNING 'ETH-CØDE'.  THE CØDESHEET
          GIVES THE CØLUMN NUMBERS ØF THE VARIØUS
          CHARACTERISTICS AND THE CØDING RELEVANT
          TØ DECISIØNS AS TØ THEIR PRESENCE ØR ABSENCE.

WHAT ARE THE CØL. NØS. ØF THE CHARACTERISTICS ? 8,59

WHAT ARE THE NAMES ØF THESE TWØ CHARACTERISTICS IN ØRDER
? FISHING,SEX DIFFERENCE

FØR FISHING TYPE THE SYMBØLS (ØMIT CØMMAS) THAT YØU
TAKE TØ INDICATE PRESENCE ØF THAT CHAR. (ØMIT SPACES)
? 3456789

FØR SEX DIFFERENCE TYPE THE SYMBØLS (ØMIT CØMMAS) THE YØU TAKE
TØ INDICATE PRESENCE ØF THAT CHAR. (ØMIT SPACES)
? 12

DØ YØU WISH TØ RUN THESE CHARACTERISTICS HØLDING ANY CØNDITIØN
CØNSTANT? TYPE YES ØR NØ!
? YES

TYPE THE CØLUMN NØS. ØF THE CHAR. (UP TØ FIVE) WITH CØMMAS
BETWEEN THEM, YØU WISH CØNSTANT.
? 1
FØR EACH ØF THESE CHARACTERISTICS TYPE THE SYMBØLS INDICATING
THE ATTRIBUTES YØU WISH TØ INCLUDE (ØMIT CØMMAS)
CHAR.  1
? 4

NØ INFØ ØN CHARACTERISTIC FISHING IN                1  CULTURES

NØ INFØ ØN CHARACTERISTIC SEX DIFFERENCE IN         397  CULTURES

CHARACTERISTIC FISHING  ABSENT IN                   295   CULTURES
CHARACTERISTIC SEX DIFFERENCE  ABSENT IN            290   CULTURES

CHARACTERISTIC FISHING(4) PRESENT IN                872   CULTURES
CHARACTERISTIC SEX DIFFERENCE(12) PRESENT IN        481   CULTURES

ØF 1168 CULTURES 659  HAVE BEEN EXCLUDED BY CØNDITIØN

                    TABLE ØF RELATIØNSHIPS
                HØLDING 1  CØNDITIØN(S) CØNSTANT

                          FISHING
                             -     +
                          ---------------
                     +       6    60
           SEX DIFFERENCE          :
                     -       5    41
                          ---------------


X2 ØN YØUR TABLE RUNNING CHARACTERISTIC ØF FISHING(4)AGAINST
CHARACTERISTIC ØF SEX DIFFERENCE(12) IS  0.0001

PRØBABILITY IS.....  0.9874

Q IS...  0.0989

TIME: 6.035 SEC.
```

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 6 seconds, at a rate of \$0.11/sec. Total terminal time was about 7 minutes, at a rate of \$3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead
= \$0.66 + \$0.41 + communications + network overhead
= \$1.07 + communications + network overhead

CONTENTS—ETHATLAS

pages

1– 2 Identification & Abstract
3 User Instructions
5 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Cultural Comparison
CALLING NAME	CULTCOMP
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Revised by Mark Hebenstreit, Class of 1970 Data from the "Ethnographic Atlas" were obtained from Herbert Barry III, Department of Anthropology, University of Pittsburgh
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

CULTCOMP retrieves coded data from George Peter Murdock's *Ethnographic Atlas* and translates them for cultural comparison. The user learns the degree of similarity and dissimilarity between the two cultures of his choice, and has the option to see a print-out in chart form which gives relevant characteristics for each culture. Data are available on 1168 cultures. In the event a codesheet is not available, the user has the option to list all 1168 cultures for examination.

REFERENCES

Murdock, G.P., *Ethnology Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

continued

Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

"Ethnographic Atlas", by the editors of *Ethnology, I*, 1,(Jan., 1962)-*VII*, 3, (July, 1968), installments of information in each issue.

Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2,(April, 1967).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook"[1]. Respond to questions as they appear. When typing in the names of two cultures be sure to separate them by a comma, e.g. AZTEC,BABYLONIAN. The user is warned that the listing of the 1168 cultues—if that option is specified—creates several pages of output and consumes considerable terminal time. It is advisable to obtain a copy of the "Ethnographic Atlas Codesheet" (Revised)[2] as indicated below.

REFERENCES

1. Morton, A. K., *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., June, 1970). Available from the EIN Office at the cost of reproduction and mailing.

2. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
ETHNOGRAPHIC ATLAS              -1-             CULTURAL COMPARISON


WOULD YOU LIKE TO SEE A LIST OF THE 1168 CULTURES? NO

DO YOU WANT TO CONSIDER 1 OR 2 CULTURES? 2

WHICH TWO CULTURES WOULD YOU LIKE COMPARED? AZTEC,BABYLONIAN

THE TWO CULTURES, AZTEC AND BABYLONIAN, ARE SIMILAR IN 38
CHARACTERISTICS, AND DISSIMILAR IN 37 CHARACTERISTICS.

WOULD YOU LIKE A TABLE PRINTED OUT COMPARING THE CHARACTERISTICS
OF EACH COLUMN FOR THE TWO CULTURES? YES

     BABYLONIAN                                  AZTEC
========================================================================
=CIRCUM-MEDITERRANEAN     *REGIONAL ID       *NORTH AMERICA            =
=0                        *  LESSER AREAS    *0                        =
=4                        *  SERIAL NUMBER   *2                        =
=                         *SUBSIST. ECON.:   *                         =
=0-5%                     *  GATHERING       *0-5%                     =
=0-5%                     *  HUNTING         *6-15%                    =
=16-25%                   *  FISHING         *16-25%                   =
=16-25%                   *  ANIMAL HUSB.    *0-5%                     =
=56-65%                   *  AGRICULTURE     *66-75%                   =
=                         *MODE OF MARRIAGE: *                         =
=TOKEN BRIDE-PRICE        *  PREVAILING      *TOKEN BRIDE-PRICE        =
=DOWRY                    *  ALTERNATE       *NO ALTERNATIVE MODE      =
=                         *FAMILY ORGAN.:    *                         =
=INDEP NUCLEAR FAMILIES   *  PREVAILING      *INDEP NUCLEAR FAMILIES   =
=  WITH MONOGAMY          *                  *  WITH LIMITED POLYGYNY  =
=POLYANDRY--NO EXTENDED   *  MARITAL COMP.   *POLYANDRY--NO EXTENDED   =
=  FAMILY                 *                  *  FAMILY                 =
=                         *MARITAL RESIDENCE:*                         =
=NO DEVIATIONS FROM PREVA-*  1 YR (IF DIFF)  *NO DEVIATIONS FROM PREVA-=
=  LENT RESIDENCE OR NON  *                  *  LENT RESIDENCE OR NON  =
=  ESTABLISHED HOUSEHOLD  *                  *  ESTABLISHED HOUSEHOLD  =
=AMBILOCAL BILOCAL OR     *  PREVALENT       *VIRILOCAL                =
=  UTROLOCAL              *                  *                         =
=NO DEVIATIONS FROM PREVA-*  ALTERNATIVES    *NO DEVIATIONS FROM PREVA-=
=  LENT RESIDENCE OR NON  *                  *  LENT RESIDENCE OR NON  =
=  ESTABLISHED HOUSEHOLD  *                  *  ESTABLISHED HOUSEHOLD  =
=                         *COMMUNITY ORGAN.: *                         =
=INSUFFICIENT INFO        *  PRIMARY         *SEGMENTED COMMUNITIES    =
=                         *                  *  WITHOUT LOCAL EXOGAMY  =
=NO SECONDAR ORGAN.       *  SECONDARY       *NO SECONDAR ORGAN.       =
=                         *PATLIN MATLIN EXO *                         =
=ABSENCE OF KIN GROUPS &  *  LARGEST PATLIN  *ABSENCE OF KIN GROUPS &  =
=  ALSO OF EXOGAMY        *                  *  ALSO OF EXOGAMY        =
=NO DIFFERENCE FROM       *  LGST PATLIN EXO *NO DIFFERENCE FROM       =
=  PRECEDING COLUMN       *                  *  PRECEDING COLUMN       =
=ABSENCE OF KIN GROUPS &  *  LARGEST MATLIN  *ABSENCE OF KIN GROUPS &  =
=  ALSO OF EXOGAMY        *                  *  ALSO OF EXOGAMY        =
=NO DIFFERENCE FROM       *  LGST MATLIN EXO *NO DIFFERENCE FROM       =
=  PRECEDING COLUMN       *                  *  PRECEDING COLUMN       =
=                         *COGNATIC KIN GPS: *                         =
=BILATERAL DESCENT        *  PRIMARY         *RAMAGES (ANCESTOR-ORIENT-=
=                         *                  *  ED AMBILINEAL KIN GROUPS=
=                         *                  *  IF NOT EXOGAMOUS)      =
=ABSENCE OF COGNATIC KIN  *  SECONDARY       *ABSENCE OF COGNATIC KIN  =
=  GROUPS (PRESENCE OF    *                  *  GROUPS (PRESENCE OF    =
=  UNILINEAL DESCENT)     *                  *  UNILINEAL DESCENT)     =
=                         *COUSIN MARRIAGE:  *                         =
=INSUFFICIENT INFO        *  TYPES ALLOWED   *NONLATERAL--EVIDENCE     =
=                         *                  *  AVAILABLE ONLY FOR     =
=                         *                  *  1ST COUSIN             =
========================================================================
```

continued

```
CULTURAL COMPARISON          -2-          CONTINUED

          BABYLONIAN                           AZTEC
========================================================================
=NO PREFERENTIAL OR PRE-  *  PREF/PRES.      *NO PREFERENTIAL OR PRE-  =
=  SCRIPTIVE UNIONS       *                  *  SCRIPTIVE UNIONS       =
=INSUFFICIENT INFO        *KIN TERM 1ST COUS *HAWAIIAN                 =
=                         *AGRICULTURE       *                         =
=INTENSIVE AGRICULTURE    *  INTEN OF CULTIVA*INTENSIVE AGRICULTURE    =
=  LARGELY DEPENDENT      *                  *  LARGELY DEPENDENT      =
=  ON IRRIGATION          *                  *  ON IRRIGATION          =
=CEREAL GRAINS            *  MAIN CROP TYPE  *CEREAL GRAINS            =
=COMPACT AND RELATIVELY   *SETTLEMENT PATTERN*COMPACT AND RELATIVELY   =
=  PERMANENT SETTLEMENTS  *                  *  PERMANENT SETTLEMENTS  =
=ONE OR MORE CITIES OF    *MEAN LOCAL COMMUN *ONE OR MORE CITIES OF    =
=  MORE THAN 50000        *                  *  MORE THAN 50000        =
=                         *JURIS HIERARCHY   *                         =
=TWO LEVELS               *  LOCAL COMMUNITY *THREE LEVELS             =
=THREE LEVELS             *  BEYOND LOCAL    *TWO LEVELS               =
=ABSENT OR NOT REPORTED   *HIGH GODS         *PRESENT AND ACTIVE IN HU-=
=                         *                  *  MAN AFFAIRS-NOT SUPPOR-=
=                         *                  *  TIVE OF HUMAN MORALITY =
=INSUFFICIENT INFO        *TYPES OF GAMES    *ALL THREE TYPES          =
=INSUFFICIENT INFO        *P-P SEX TABOOS    *INSUFFICIENT INFO        =
=INSUFFICIENT INFO        *MALE GENITAL MUT  *ABSENT OR NOT GENERALLY  =
=                         *                  *  PRACTICED              =
=INSUFFICIENT INFO        *SEG OF ADOL BOYS  *INSUFFICIENT INFO        =
=                         *TYPE ANIMAL HUSB.:*                         =
=PRIOR TO CONTACT PERIOD  *  PLOW CULTIVATION*NO PLOW ANIMALS OR INSUF-=
=                         *                  *  FICIENT INFORMATION    =
=BOVINE ANIMALS (CATTLE   *  PREDOMINANT TYPE*ABSENCE OR NEAR ABSENCE  =
=  MITHUN WATER BUFFALO   *                  *  OF LARGE DOMESTIC      =
=  YAKS)                  *                  *  ANIMALS                =
=MILKED MORE OFTEN THAN   *  MILKING DOMS ANI*LITTLE OR NO MILKING OR  =
=  SPORADICALLY           *                  *  INSUFFICIENT INFOR-    =
=                         *                  *  MATION                 =
=INTENSIVE AGRI (NOTE     *SUM-SUBSIS ECON   *INTENSIVE AGRI (NOTE     =
=  ABOVE) > ANIMAL        *                  *  ABOVE) > ANIMAL        =
=  HUSBANDRY              *                  *  HUSBANDRY              =
=BILATERAL                *SUM-DESCENT       *AMBILINEAL               =
=                         *METAL WORKING:    *                         =
=MALES ALONE OR ALMOST    *  SEX DIFFERENCE  *MALES ALONE OR ALMOST    =
=  ALONE                  *                  *  ALONE                  =
=CRAFT SPECIALIZATION     *  SPECIALIZATION  *CRAFT SPECIALIZATION     =
=                         *WEAVING:          *                         =
=ACTIVITY IS PRESENT BUT  *  SEX DIFFERENCE  *FEMALES ALONE OR ALMOST  =
=  SEX PARTICIPATION      *                  *  ALONE                  =
=  IS UNSPECIFIED         *                  *                         =
=ACTIVITY ABSENT OR INSUF-*  SPECIALIZATION  *ACTIVITY ABSENT OR INSUF-=
=  FICIENT INFORMATION    *                  *  FICIENT INFORMATION    =
=                         *LEATHER WORKING:  *                         =
=ACTIVITY IS PRESENT BUT  *  SEX DIFFERENCE  *MALES ALONE OR ALMOST    =
=  SEX PARTICIPATION      *                  *  ALONE                  =
=  IS UNSPECIFIED         *                  *                         =
=ACTIVITY ABSENT OR INSUF-*  SPECIALIZATION  *CRAFT SPECIALIZATION     =
=  FICIENT INFORMATION    *                  *                         =
=                         *POTTERY:          *                         =
=ACTIVITY IS PRESENT BUT  *  SEX DIFFERENCE  *MALES ALONE OR ALMOST    =
=  SEX PARTICIPATION      *                  *  ALONE                  =
=  IS UNSPECIFIED         *                  *                         =
=ACTIVITY ABSENT OR INSUF-*  SPECIALIZATION  *CRAFT SPECIALIZATION     =
=  FICIENT INFORMATION    *                  *                         =
=                         *BOAT BUILDING:    *                         =
=ACTIVITY IS PRESENT BUT  *  SEX DIFFERENCE  *ACTIVITY IS PRESENT BUT  =
=  SEX PARTICIPATION      *                  *  SEX PARTICIPATION      =
=  IS UNSPECIFIED         *                  *  IS UNSPECIFIED         =
=ACTIVITY ABSENT OR INSUF-*  SPECIALIZATION  *ACTIVITY ABSENT OR INSUF-=
=  FICIENT INFORMATION    *                  *  FICIENT INFORMATION    =
=                         *HOUSE CONSTRUCTION*                         =
=INSUFFICIENT INFO        *  SEX DIFFERENCE  *MALES ALONE OR ALMOST    =
=                         *                  *  ALONE                  =
=ACTIVITY ABSENT OR INSUF-*  SPECIALIZATION  *CRAFT SPECIALIZATION     =
=  FICIENT INFORMATION    *                  *                         =
========================================================================
```

continued

```
CULTURAL COMPARISON             -3-             CONTINUED

           BABYLONIAN                                  AZTEC
****************************************************************************
*                          *GATHERING:          *                          *
*INSUFFICIENT INFO         *  SEX DIFFERENCE    *ACTIVITY IS ABSENT OR     *
*                          *                    *  UNIMPORTANT             *
*ACTIVITY ABSENT OR INSUF- *  SPECIALIZATION    *ACTIVITY ABSENT OR INSUF- *
*  FICIENT INFORMATION     *                    *  FICIENT INFORMATION     *
*                          *HUNTING:            *                          *
*INSUFFICIENT INFO         *  SEX DIFFERENCE    *MALES ALONE OR ALMOST     *
*                          *                    *  ALONE                   *
*ACTIVITY ABSENT OR INSUF- *  SPECIALIZATION    *ACTIVITY ABSENT OR INSUF- *
*  FICIENT INFORMATION     *                    *  FICIENT INFORMATION     *
*                          *FISHING:            *                          *
*INSUFFICIENT INFO         *  SEX DIFFERENCE    *MALES ALONE OR ALMOST     *
*                          *                    *  ALONE                   *
*ACTIVITY ABSENT OR INSUF- *  SPECIALIZATION    *CRAFT SPECIALIZATION      *
*  FICIENT INFORMATION     *                    *                          *
*                          *ANIMAL HUSBANDRY:   *                          *
*MALES ALONE OR ALMOST     *  SEX DIFFERENCE    *ACTIVITY IS ABSENT OR     *
*  ALONE                   *                    *  UNIMPORTANT             *
*ACTIVITY ABSENT OR INSUF- *  SPECIALIZATION    *ACTIVITY ABSENT OR INSUF- *
*  FICIENT INFORMATION     *                    *  FICIENT INFORMATION     *
*                          *AGRICULTURE:        *                          *
*MALES ALONE OR ALMOST     *  SEX DIFFERENCE    *MALES APPRECIABLY MORE    *
*  ALONE                   *                    *                          *
*                          *LINGUISTIC AFFIL.   *                          *
*1                         *  PHYLUM            *8                         *
*AFRO-ASIATIC              *  FAMILY            *NAHUATLAN                 *
*6                         *  SUBFAMILY         *1                         *
*                          *CLASS STRAT.        *                          *
*COMPLEX (SOCIAL CLASSES)  *  PREVAILING        *COMPLEX (SOCIAL CLASSES)  *
*NO ADDITIONAL IMPORTANT   *  ADD. FEATURE      *NO ADDITIONAL IMPORTANT   *
*  FEATURE OR ABSENCE OF   *                    *  FEATURE OR ABSENCE OF   *
*  STRATIFICATION          *                    *  STRATIFICATION          *
*                          *CASTE STRAT.        *                          *
*ABSENT OR INSIGNIFICANT   *  PREVAILING        *ABSENT OR INSIGNIFICANT   *
*NO SECOND TYPE OR ABSENCE *  SECOND TYPE       *NO SECOND TYPE OR ABSENCE *
*  OF STRATIFICATION       *                    *  OF STRATIFICATION       *
*                          *SLAVERY             *                          *
*HEREDITARY AND SOCIALLY   *  TYPE              *INCIPIENT OR NONHEREDI-   *
*  SIGNIFICANT             *                    *  TARY                    *
*ABSENT OR NO DIFFERENCE   *  FORMER PRESENCE   *ABSENT OR NO DIFFERENCE   *
*  FROM PRECEDING COL.     *                    *  FROM PRECEDING COL.     *
*INSUFFICIENT INFO         *LOCAL HEADMAN       *HEREDITARY BY A SON       *
*                          *INHER. REAL PROP.   *                          *
*PATRILINEAL (SONS)        *  RULE OR PRACTICE  *PATRILINEAL (SONS)        *
*EQUAL OR RELATIVELY EQUAL *  DISTRIBUTION      *PRIMOGENITURE (TO THE     *
*                          *                    *  SENIOR INDIVIDUAL)      *
*                          *INHER MOVABLE PROP  *                          *
*PATRILINEAL (SONS)        *  RULE OR PRACTICE  *PATRILINEAL (SONS)        *
*EQUAL OR RELATIVELY EQUAL *  DISTRIBUTION      *PRIMOGENITURE (TO THE     *
*                          *                    *  SENIOR INDIVIDUAL)      *
*INSISTENCE ON VIRGINITY   *PREMAR-SEX-GIRLS    *PROHIBITED BUT WEAKLY     *
*                          *                    *  CENSURED & NOT          *
*                          *                    *  INFREQUENT              *
*                          *GROUND PLAN DWELL   *                          *
*RECTANGULAR OR SQUARE     *  PREVAILING        *RECTANGULAR OR SQUARE     *
*NO SECONDARY TYPE         *  SECONDARY         *QUADRANGULAR AROUND (OR   *
*                          *                    *  PARTIALLY AROUND) AN    *
*                          *                    *  INNER COURT             *
*                          *FLOOR LEVEL         *                          *
*FLOOR FORMED BY GROUND    *  PREVAILING        *ELEVATED SLIGHTLY OR ON   *
*                          *                    *  RAISED PLATFORM         *
*NO SECONDARY TYPE         *  SECONDARY         *FLOOR FORMED BY GROUND    *
*                          *WALL MATERIAL       *                          *
*ADOBE CLAY OR             *  PREVAILING        *ADOBE CLAY OR             *
*  DRIED BRICK             *                    *  DRIED BRICK             *
*NO SECONDARY TYPE         *  SECONDARY         *STONE STUCCO CONCRETE     *
*                          *                    *  OR FIRED BRICK          *
****************************************************************************
```

continued

```
CULTURAL COMPARISON              -4-              CONTINUED

          BABYLONIAN                              AZTEC
======================================================================
=                       *SHAPE OF ROOF     *                         =
=FLAT OR HORIZONTAL     *  PREVAILING      *FLAT OR HORIZONTAL       =
=NO SECONDARY TYPE      *  SECONDARY       *FLAT OR HORIZONTAL       =
=                       *ROOF MATERIAL     *                         =
=EARTH OR TURF          *  PREVAILING      *GRASS LEAVES BRUSH       =
=                       *                  *  OR OTHER THATCH        =
=NO SECONDARY TYPE      *  SECONDARY       *NO SECONDARY TYPE        =
=INSUFFICIENT INFO      *POLITICAL INTEGRAT*STATES (AT LEAST         =
=                       *                  *  100,000)               =
=INSUFFICIENT INFO      *POLITICAL SUCCESS.*PATRILINEAL A SON PREFER-=
=                       *                  *  RED TO A YOUNGER       =
=                       *                  *  BROTHER                =
=                       *ENVIRONMENT       *                         =
=INSUFFICIENT INFO      *  PRIMARY         *TEMPERATE FOREST (MOSTLY =
=                       *                  *  MOUNTAINOUS)           =
=INSUFFICIENT INFO      *  SECONDARY       *NO SECONDARY ENVIRONMENT =
======================================================================


REFERENCE MATERIAL
------------------

AZTEC

NO REFERENCES

BABYLONIAN

NO REFERENCES





----------------------------------------------------------------------

TIME: 6.663 SEC.
```

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 6.66 seconds, at a rate of $0.11/sec. Total terminal time was about 30 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead

= $0.73 + $1.75 + communications + network overhead

= $2.48 + communications + network overhead

CONTENTS—CULTCOMP

pages
1– 2 Identification & Abstract
3 User Instructions
5– 8 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Culture Identification
CALLING NAME	CULTPIK
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Program written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE-635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

CULTPIK allows the user to identify those cultures which satisfy any selected set of characteristics. Characteristics may be selected from Murdock's *Ethnographic Atlas*[1] or from Textor's *Cross-Cultural Summary*[2]. Complete codesheets may be obtained from the Department of Anthropology, College Museum, Dartmouth College. Portions of the codesheets may be examined by running either ETH—CODE (EIN No. 000 0119) or TEX—CODE (EIN No. 000 0127).

REFERENCES

1. Murdock, G.P., *Ethnographic Atlas*, (Univ. of Pittsburgh Press, Pittsburgh, Pa., 1967).

2. Textor, R.B., Comp., *A Cross Cultural Summary*, (HRAF Press, New Haven, Conn., 1967).

3. "Cross-Cultural Survey", Codesheet for R.B. Textor's finished characteristics from *A Cross-Cultural Summary*, (Dept. of Anthrop., College Museum, Dartmouth College, Hanover, N.H.). Available from the EIN Office at the cost of reproduction and mailing.

continued

4. Barry III, H., "Ethno", (*Ethnology* version of *Ethnographic Atlas*, Univ. of Pittsburgh, Pittsburgh, Pa., 1970).

5. "Ethnographic Atlas" by the editors of *Ethnology, I*, 1,(Jan., 1962)—*VII*, 3,(July, 1968), installments of information in each issue.

6. Murdock, G.P., "Ethnographic Atlas: A Summary", *Ethnology, VI*, 2, (April, 1967).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook"[1]. Ask to see the instructions, which specify the format in which the choice of characteristics must be entered. Supply the options as requested.

The user may specify as many as 15 characteristics for either set of data.

REFERENCES

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., June, 1970). Available from the EIN Office at the cost of reproduction and mailing.

2. "Cross Cultural Survey", Codesheet for R.B. Textor's finished characteristics from *A Cross-Cultural Summary*, (Dept. of Anthrop., College Museum, Dartmouth College, Hanover, N.H.). Available from the EIN Office at the cost of reproduction and mailing.

3. "Ethnographic Atlas Codesheet", (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H., 1970), Revised. Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
CULTPIK    21 SEP 70  15:58

     WOULD YOU LIKE TO SEE THE INSTRUCTIONS FOR THIS PROGRAM? YES
-----------------------------------------------------------------



                         ==========================
                         ==CULTPIK  INSTRUCTIONS==
                         ==========================

         THIS PROGRAM ALLOWS THE USER TO IDENTIFY THOSE CULTURES
     WHICH SATISFY  ANY  SELECTED  SET  OF  CHARACTERISTICS.  THE
     CULTURES WHICH  HAVE  THESE CHARACTERISTICS MAY BE  SELECTED
     FROM  EITHER  THE  DATA  COLLECTED  BY  GEORGE PETER MURDOCK
     (ETHNOGRAPHIC ATLAS), OR ROBERT B.  TEXTOR (A CROSS-CULTURAL
     SURVEY). CODESHEETS FOR THESE TWO SETS OF DATA CAN BE LISTED
     DURING THE RUN OF THIS PROGRAM AND  ARE  NECESSARY  FOR  THE
     SPECIFICATION OF CHARACTERISTICS.  THE  COMPLETE  CODESHEETS
     CAN BE OBTAINED  FROM  THE OFFICE  OF ANTHROPOLOGY,  COLLEGE
     MUSEUM,  DARTMOUTH  COLLEGE.  THE  CODESHEETS  CAN  ALSO  BE
     OBTAINED BY RUNNING EITHER ETH-CODE*** OR TEX-CODE***.

         YOU MAY SPECIFY  AS  MANY  AS  15  CHARACTERISTICS  FOR
     EITHER SET OF DATA. CHARACTERISTICS MUST BE SPECIFIED IN THE
     FOLLOWING FORMAT, TYPING THE INFO FOR EACH CHARACTERISTIC ON
     A SEPARATE LINE --  WHEN  YOU HAVE SUPPLIED THE INFO FOR ALL
     YOUR CHARACTERISTICS, PRESS THE CARRIAGE RETURN KEY (CR).


          MURDOCK DATA:

               FIRST  TYPE  THE  COLUMN  NUMBER  (WHICH
               APPEARS  ON  THE   FAR   LEFT   OF   THE
               CODESHEET) AND  THEN  THOSE CODES  WHICH
               INDICATE THE  PRESENCE  OF  A TRAIT  YOU
               WISH TO INCLUDE  IN YOUR SELECTION. SOME
               CHARACTERISTICS  DO NOT HAVE THEIR CODES
               SPECIFIED  IN THE CODESHEET (FOR EXAMPLE
               'LESSER     AREAS'     AND     'LINGUISTIC
               AFFILIATION'), BUT APPROPRIATE CODES MAY
               BE TYPED IN.

               FOR EXAMPLE: ? 94, 7, 8, 5 (NOTE USE OF COMMAS!)


          TEXTOR DATA:

               FIRST TYPE THE  NUMBER OF THE  FINISHED
               CHARACTERISTIC  (AS IT APPEARS  ON  THE
               CODESHEET), AND THEN ONE OR MORE OF THE
               FOLLOWING CODES:
                              1.) LEFT SENTENCE
                              2.) RIGHT SENTENCE
                              3.) AMBIGUOUS
                              4.) IRRELEVANT
                              5.) UNASCERTAINED

          FOR EXAMPLE: ? 278, 1

      YOU HAVE  THE  OPTION TO HAVE ALL THOSE CULTURES LISTED
 WHICH SATISFY ANY  CHARACTERISTIC,  OR  ONLY  THOSE CULTURES
 WHICH SATISFY ALL SPECIFICATIONS.

      FOR INFORMATION ON OTHER PROGRAMS RELATED TO THIS, LIST
 ANTHRO***.
```

continued

```
ARE YOU INTERESTED IN THE MURDØCK ØR TEXTØR DATA? MURDØCK

INDICATE THE NUMBER ØF YØUR CHØICE: 1.) CULTURES WHICH
      SATISFY ALL ØF THE SPECIFICATIØNS, 2.) CULTURES WHICH
      SATISFY ANY ØF THE SPECIFICATIØNS (CØLUMN NUMBERS ARE
      PRINTED WITH THE CULTURES)     ? 1

DØ YØU WANT TØ SEE THE CØDESHEET FØR THE MURDØCK DATA? NØ

TYPE IN THE CØLUMN NUMBERS AND CØDES FØR THE MURDØCK DATA
      ACCØRDING TØ THE DIRECTIØNS. IF YØU HAVE LESS THAN 15
      CØLUMNS, PRESS THE CARRIAGE RETURN KEY AFTER THE LAST
      ENTRY.

? 1,4
? 8,5
? 59,1
? 67,1
?



 2 SELECTED CULTURES FRØM MURDØCK DATA

          CØDES 4 INDICATE PRESENCE FØR CØLUMN 1.
          CØDES 5 INDICATE PRESENCE FØR CØLUMN 8.
          CØDES 1 INDICATE PRESENCE FØR CØLUMN 59.
          CØDES 1 INDICATE PRESENCE FØR CØLUMN 67.

                          ---<>---

CARØLINIAN        ØNTØNG-JAV




--------------------------------------------------

TIME: 9.207 SEC.
```

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 9.2 seconds, at a rate of $0.11/sec. Total terminal time was about 10 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead

= $1.01 + $0.58 + communications + network overhead

= $1.59 + communications + network overhead

CONTENTS—CULTPIK

pages

1– 2 Identification & Abstract
3 User Instructions
5– 6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Finished Characteristics of 400 Cultures
CALLING NAME	TEXTOR
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College Data for the 400 cultures were compiled by Robert B. Textor
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646-2864

FUNCTIONAL ABSTRACT

TEXTOR reproduces the dichotomies for the finished characteristics of 400 cultures as constructed by Robert B. Textor[1]. The user has the option of seeing a list of the major characteristic categories, as well as selected sub-divisions, just as in TEX–CODE (EIN No. 000 0127).

REFERENCES

1. Textor, R.B., Comp., *A Cross-Cultural Summary*, (HRAF Press, New Haven, Conn., 1967).

450

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook"[1]. The program will ask for the numbers of the two finished characteristics in which you are interested. These numbers should be separated by a comma, e.g. 3,17. The first finished characteristic will become the subject of the two sentences to be produced, and the second will be the object.

REFERENCES

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., 1970). Available from the EIN Office at the cost of reproduction and mailing.

2. "Cross-Cultural Survey", Codesheet for R.B. Textor's finished characteristics from *A Cross-Cultural Summary*, (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H.). Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
                                 -1-
ROBERT B. TEXTOR'S CROSS CULTURAL SURVEY                    10/09/70

WOULD YOU LIKE TO SEE THE CODESHEET FOR THE FINISHED
   CHARACTERISTICS? YES

 1 LOCATION                         2 LINGUISTIC AFFILIATION
 3 NATURAL ENVIRONMENT              4 SETTLEMENT PATTERN
 5 DIET                             6 SUBSISTENCE
 7 TECHNOLOGY                       8 WRITING SYSTEM
 9 DEMOGRAPHY                       10 POLITICAL ORGANIZATION
 11 SOCIETAL COMPLEXITY             12 STRATIFICATION
 13 WORK ORGANIZATION               14 OCCUPATIONAL SPECIALIZATION
 15 ECONOMICS                       16 JUSTICE AND LAW
 17 JURISPRUDENCE & MEDICINE        18 COMMUNITY ORGANIZATION
 19 LARGEST NONCOGNATIC KIN GROUP   20 LINEALITY OF KIN GROUP
 21 INHERITANCE                     22 MARITAL RESIDENCE
 23 COUSIN MARRIAGE                 24 COUSIN TERMINOLOGY
 25 FAMILY ORGANIZATION             26 POLYGYNY
 27 AUTHORITY WITHIN FAMILY         28 AVOIDANCE
 29 MODE OF MARRIAGE                30 DIVORCE
 31 STATUS OF WOMEN                 32 FERTILITY
 33 PREGNANCY & CHILDBIRTH          34 INFANCY & CHILDHOOD
 35 ADOLESCENCE                     36 SEXUAL BEHAVIOR
 37 ILLNESS AND THERAPY             38 AGGRESSION AND WARFARE
 39 RELIGION MAGIC AND ESCHATOLOGY  40 GAMES
 41 CULTURE CONTACT & CHANGE        42 MISCELLANEOUS
 43 METHODOLOGICAL SECTION          44 'WHISKERS' CHARACTERISTICS

BE SELECTIVE AND DECIDE ON THE MINIMUM NUMBER OF FINISHED
   CHARACTERISTICS WHICH INTEREST YOU, BECAUSE THE PRINT-OUT
   IS VERYY LONG. THE COMPLETE CODESHEET IS AVAILABLE AT THE
   ANTHROPOLOGY OFFICE IN WILSON HALL.

FOR WHICH TITLE DIVISIONS WOULD YOU LIKE TO SEE THE
   FINISHED CHARACTERISTICS? 2,8




CODESHEET FOR TEXTOR'S CROSS CULTURAL SURVEY

                        --LINGUISTIC AFFILIATION--

   FINISHED CHARACTERISTIC  17          CULTURES WHOSE . . .
(LEFT) LINGUISTIC AFFILIATION IS NIGER-CONGO (61).
(RIGHT) LINGUISTIC AFFILIATION IS OTHER THAN NIGER-CONGO (339).

   FINISHED CHARACTERISTIC  18          CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS CHARI-NILE (14).
(R) LINGUISTIC AFFILIATION IS OTHER THAN CHARI-NILE (386).

   FINISHED CHARACTERISTIC  19          CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS AFRO-ASIATIC (21).
(R) LINGUISTIC AFFILIATION IS OTHER THAN AFRO-ASIATIC (379).

   FINISHED CHARACTERISTIC  20          CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS NIGER-CONGO, RATHER THAN AFRO-
   ASIATIC (61).
(R) LINGUISTIC AFFILIATION IS AFRO-ASIATIC, RATHER THAN NIGER-
   CONGO (21).
```

continued

-2-

TEXTOR'S CODESHEET (CONT.)

FINISHED CHARACTERISTIC 21 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS INDO-EUROPEAN (28).
(R) LINGUISTIC AFFILIATION IS OTHER THAN INDO-EUROPEAN (372).

FINISHED CHARACTERISTIC 22 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS INDO-EUROPEAN, RATHER THAN AFRO-ASIATIC (28).
(R) LINGUISTIC AFFILIATION IS AFRO-ASIATIC, RATHER THAN INDO-EUROPEAN (21).

FINISHED CHARACTERISTIC 23 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS ALTAIC (11).
(R) LINGUISTIC AFFILIATION IS OTHER THAN ALTAICC (389).

FINISHED CHARACTERISTIC 24 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS MALAYO-POLYNESIAN (60).
(R) LINGUISTIC AFFILIATION IS OTHER THAN MALAYO-POLYNESIAN (340).

FINISHED CHARACTERISTIC 25 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS TIBETO-BURMAN (12).
(R) LINGUISTIC AFFILIATION IS OTHER THAN TIBETO-BURMAN (388).

FINISHED CHARACTERISTIC 26 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS MON-KHMER (10).
(R) LINGUISTIC AFFILIATION IS OTHER THAN MON-KHMER (390).

FINISHED CHARACTERISTIC 27 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS MALAYO-POLYNESIAN, RATHER THAN TIBETO-BURMAN (60).
(R) LINGUISTIC AFFILIATION IS TIBETO-BURMAN, RATHER THAN MALAYO-POLYNESIAN (12).

FINISHED CHARACTERISTIC 28 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS MALAYO-POLYNESIAN, RATHER THAN MON-KHMER (60).
(R) LINGUISTIC AFFILIATION IS MON-KHMER, RATHER THAN MALAYO-POLYNESIAN (10).

FINISHED CHARACTERISTIC 29 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS TIBETO-BURMAN, RATHER THAN MON-KHMER (12).
(R) LINGUISTIC AFFILIATION IS MON-KHMER, RATHER THAN TIBETO-BURMAN (10).

FINISHED CHARACTERISTIC 30 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS ATHABASKAN (8).
(R) LINGUISTIC AFFILIATION IS OTHER THAN ATHABASKAN (392).

FINISHED CHARACTERISTIC 31 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS ALGONKIAN (9).
(R) LINGUISTIC AFFILIATION IS OTHER THAN ALGONKIAN (391).

FINISHED CHARACTERISTIC 32 CULTURES WHOSE . . .
(L) LINGUISTIC AFFILIATION IS CARIBAN (8).
(R) LINGUISTIC AFFILIATION IS OTHER THAN CARIBAN (392).

--WRITING SYSTEM--

FINISHED CHARACTERISTIC 77 CULTURES WHERE . . .
(L) THE WRITING SYSTEM IS ALPHABETIC-OR-PHONETIC, RATHER THAN BEING MNEMONIC OR ABSENT (15)..
(R) THE WRITING SYSTEM IS MNEMONIC OR ABSENT, RATHER THAN BEING ALPHABEETIC-OR-PHONETIC (39).

FINISHED CHARACTERISTIC 78 CULTURES WHERE . . .
(L) THE WRITING SYSTEM IS ALPHABETIC-OR-PHONETIC, OR MNEMONIC, RATHER THAN BEING ABSENT (36).
(R) A WRITING SYSTEM IS ABSENT, RATHER THAN BEING PRESENT IN EITHER ALPHABETIC-OR-PHONETIC FORM OR MNEMONIC FORM (18).

continued

```
TEXTOR (CONT.)

WHAT ARE THE COLUMN NUMBERS OF THE FINISHED CHARACTERISTICS
YOU WISH FOR YOUR SUBJECT AND THHOSE YOU WISH FOR YOUR OBJECT
? 3,17


          --SUBJECT--             .         --OBJECT--
FINISHED CHARACTERISTIC # 3      / \    FINISHED CHARACTERISTIC # 17
                                /   \
                         LEFT  / (A) \  LEFT
                              /       \
                             /    60   \
                            /\         /\
                           /  \       /  \
                          /(B) \     / (C)\
                         /      \   /      \
                        /        \ /        \
                       .    1     .    20    .
                        \        / \        /
                         \      /   \      /
                          \    / (D) \    /
                           \  /       \  /
                            \/    319  \/
                             \         /
                              \       /
                        RIGHT  \     /  RIGHT
                                \   /
                                 \./
                                  .


CHI SQUARE =  270.479

CHI-SQUARE PROBABILITY =  0

PHII COEFFICIENT = -0.831004   N =  400



LEFT SENTENCE --                                              75 % OF 80

(80) CULTURES LOCATED
IN AFRICA
TEND TO BE THOSE WHOSE
LINGUISTIC AFFILIATION IS NIGER-CONGO (61).

RIGHT SENTENCE --                                             99 % OF 320

(320) CULTURES LOCATED
OUTSIDE OF AFRICA
TEND TO BE THOSE WHOSE
LINGUISTIC AFFILIATION IS OTHER THAN NIGER-CONGO (339).

DO YOU WANT A LIST OF THE CULTURES? YES

TYPE IN THE LETTERS A,B,C, OR D SIGNIFYING THE CELL OR CELLS
   IN WHICH YOU ARE INTERESTED (NO COMMAS, E.G 'AB')? ABC
```

continued

TEXTOR (CONT.)

CELL A . . .

AMBA
ASHANTI
AZANDE
BABWA
BAJUN
BAMBARA
BAMILEKE
BANDA
BAYA
BEMBA
BETE
BIRIFOR
BOZO
GHAGGA
DOGON
FANG
FON
FUTAJALONKE
GANDA
GISU
GURE
HEHE
HERERO
ILA
JUKUN
KATAB
KIKUYU
KISSI
KPE
KUBA
LAMBA
LOZI
LUBA
MAMBILA
MBUGWE
MBUNDU
MBUTI
MENDE
MONGO
MOSSI
NDEMBU
NGONI
NUPE
NYAKYUSA
NYANEKA
NYORO
PENDE
RUNDI
SOTHO
SWAZI
TALLENSI
TENDA
THONGA
TIV
VENDA
WUTE
YAKO
YAO
YOMBE
YORUBA

CELL B . . .

WOLOF

CELL C . . .

BARI
BERGDAMA
DILLING
DOROBO
INGASSANA
KUNG
LANGO
LUO
MAMVU
MARGI
MASAI
MATAKAM
NAMA
NANDI
NUER
NYARO
SANDAWE
SHILLUK
TESO
TURKANA

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 8.34 seconds, at a rate of $0.11/sec. Total terminal time was about 30 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead

= $0.92 + $1.75 + communications + network overhead

= $2.67 + communications + network overhead

CONTENTS—TEXTOR

pages

pages	
1	Identification & Abstract
3	User Instructions
5– 8	I/O
9	Cost—Contents

DESCRIPTIVE TITLE	Finished characteristics from Textor's *A Cross-Cultural Summary*
CALLING NAME	TEX-CODE
INSTALLATION NAME	Dartmouth College Kiewit Computation Center
AUTHOR(S) AND AFFILIATION(S)	Written by William Koenig, Class of 1970, for James Fernandez, Department of Anthropology, Dartmouth College
LANGUAGE	Dartmouth BASIC
COMPUTER	GE–635
PROGRAM AVAILABILITY	Magnetic tape and listings presently available
CONTACT	A. Kent Morton, EIN Technical Representative, Kiewit Computation Center, Dartmouth College, Hanover, N.H. 03755 Tel.: (603) 646–2864

FUNCTIONAL ABSTRACT

Forty-four major divisions of Textor's finished characteristics[1] are first printed out, after which the user is given the chance to see selected portions in detail.

TEX–CODE may be used in conjunction with TEXTOR, (EIN No. 000 0126), or CULTPIK, (EIN No. 000 0125).

REFERENCES

1. Textor, R.B., Comp., *A Cross-Cultural Summary*, (HRAF Press, New Haven, Conn., 1967).

USER INSTRUCTIONS

Call the program following the procedures outlined in the "Concise Handbook"[1]. Input numbers as requested. In the event there are no categories for which you wish to see the subdivisions, simply push the carriage return. The user is cautioned that some divisions may require as many as five pages for output, and that it is advisable to obtain a complete copy of the 60-page codesheet as listed in the reference section[2]. This program should be used only if the codesheet is not available.

REFERENCES

1. Morton, A. Kent, *Concise Handbook to the Dartmouth Time-Sharing System*, (Kiewit Comput. Center, Hanover, N.H., June, 1970). Available from the EIN Office at the cost of reproduction and mailing.

2. "Cross-Cultural Survey", Codesheet for R.B. Textor's finished characteristics from *A Cross-Cultural Summary*, (Depart. of Anthrop., College Museum, Dartmouth College, Hanover, N.H.). Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT AND OUTPUT

```
                              -1-
TEXTOR CODESHEET

BELOW ARE THE FORTY-FOUR DIVISIONS UNDER WHICH THE 536
   FINISHED CHARACTERISTICS ARE GROUPED.

1 LOCATION                          2 LINGUISTIC AFFILIATION
3 NATURAL ENVIRONMENT               4 SETTLEMENT PATTERN
5 DIET                              6 SUBSISTENCE
7 TECHNOLOGY                        8 WRITING SYSTEM
9 DEMOGRAPHY                        10 POLITICAL ORGANIZATION
11 SOCIETAL COMPLEXITY              12 STRATIFICATION
13 WORK ORGANIZATION                14 OCCUPATIONAL SPECIALIZATION
15 ECONOMICS                        16 JUSTICE AND LAW
17 JURISPRUDENCE & MEDICINE         18 COMMUNITY ORGANIZATION
19 LARGEST NONCOGNATIC KIN GROUP    20 LINEALITY OF KIN GROUP
21 INHERITANCE                      22 MARITAL RESIDENCE
23 COUSIN MARRIAGE                  24 COUSIN TERMINOLOGY
25 FAMILY ORGANIZATION              26 POLYGYNY
27 AUTHORITY WITHIN FAMILY          28 AVOIDANCE
29 MODE OF MARRIAGE                 30 DIVORCE
31 STATUS OF WOMEN                   32 FERTILITY
33 PREGNANCY & CHILDBIRTH           34 INFANCY & CHILDHOOD
35 ADOLESCENCE                      36 SEXUAL BEHAVIOR
37 ILLNESS AND THERAPY              38 AGGRESSION AND WARFARE
39 RELIGION MAGIC AND ESCHATOLOGY   40 GAMES
41 CULTURE CONTACT & CHANGE         42 MISCELLANEOUS
43 METHODOLOGICAL SECTION           44 'WHISKERS' CHARACTERISTICS

BE SELECTIVE AND DECIDE ON THE MINIMUM NUMBER OF FINISHED
   CHARACTERISTICS WHICH INTEREST YOU, BECAUSE THE PRINT-OUT
   IS VERY LONG. THE COMPLETE CODESHEET IS AVAILABLE AT THE
   ANTHROPOLOGY OFFICE IN WILSON HALL.

FOR WHICH TITLE DIVISIONS WOULD YOU LIKE TO SEE THE
   FINISHED CHARACTERISTICS? 3,7
```

```
CODESHEET FOR TEXTOR'S CROSS CULTURAL SURVEY

                     --NATURAL ENVIRONMENT--

    FINISHED CHARACTERISTIC  33        CULTURES WHERE . . .
(LEFT) THE NATURAL ENVIRONMENT IS 'VERY HARSH,' I.E., DESERT,
    DESERT GRASSES AND SHRUBS, TUNDRA, OR HIGH PLATEAU STEPPE (59).
(RIGHT) THE NATURAL ENVIRONMENT IS OTHER THAN 'VERY HARSH,' I.E.,
    DESERT, DESERT GRASSES AND SHRUBS, TUNDRA, OR HIGH PLATEAU
    STEPPE (341).

    FINISHED CHARACTERISTIC  34        CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS DESERT OR DESERT GRASSES AND SHRUBS
    (40).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN DESERT OR DESERT GRASSES
    AND SHRUBS (360).

    FINISHED CHARACTERISTIC  35        CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS TUNDRA (11).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN TUNDRA (389).
```

continued

TEXTOR'S CODESHEET (CONT.)

FINISHED CHARACTERISTIC 36 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS 'VERY HARSH,' OR SUB-TROPICAL BUSH, OR TEMPERATE GRASSLAND (108).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN 'VERY HARSH,' OR SUB-TROPICAL BUSH, OR TEMPERATE GRASSLAND (292).

FINISHED CHARACTERISTIC 37 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS SUB-TROPICAL BUSH (24).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN SUB-TROPICAL BUSH (376).

FINISHED CHARACTERISTIC 38 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS TEMPERATE GRASSLAND (25).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN TEMPERATE GRASSLAND (375).

FINISHED CHARACTERISTIC 39 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS NORTHERN CONIFEROUS FOREST (21).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN NORTHERN CONIFEROUS FOREST (379).

FINISHED CHARACTERISTIC 40 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS TEMPERATE WOODLAND OR FOREST (35).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN TEMPERATE WOODLAND OR FOREST (365).

FINISHED CHARACTERISTIC 41 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS TROPICAL GRASSLAND (64).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN TROPICAL GRASSLAND (336).

FINISHED CHARACTERISTIC 42 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS TROPICAL OR SUB-TROPICAL RAIN FOREST, OR MONSOON FOREST (156).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN TROPICAL OR SUB-TROPICAL RAIN FOREST, OR MONSOON FOREST (244).

FINISHED CHARACTERISTIC 43 CULTURES WHERE . . .
(L) THE NATURAL ENVIRONMENT IS TROPICAL RAIN FOREST (115).
(R) THE NATURAL ENVIRONMENT IS OTHER THAN TROPICAL RAIN FOREST (285).

--TECHNOLOGY--

FINISHED CHARACTERISTIC 71 CULTURES WHERE . . .
(L) METAL WORKING IS PRESENT (98).
(R) METAL WORKING IS ABSENT (153).

FINISHED CHARACTERISTIC 72 CULTURES WHERE . . .
(L) THE PLOW IS PRESENT (61).
(R) THE PLOW IS ABSENT (339).

FINISHED CHARACTERISTIC 73 CULTURES WHERE . . .
(L) WEAVING IS PRESENT (118).
(R) WEAVING IS ABSENT (130).

FINISHED CHARACTERISTIC 74 CULTURES WHERE . . .
(L) POTTERY IS PRESENT (145).
(R) POTTERY IS ABSENT (93).

FINISHED CHARACTERISTIC 75 CULTURES WHERE . . .
(L) LEATHER WORKING IS PRESENT (123).
(R) LEATHER WORKING IS ABSENT (53).

FINISHED CHARACTERISTIC 76 CULTURES WHERE . . .
(L) BOAT BUILDING IS PRESENT (142).
(R) BOAT BUILDING IS ABSENT (63).

COST ESTIMATE

For the job included as the Sample Output, the total central processor unit (CPU) time was 1.52 seconds, at a rate of $0.11/sec. Total terminal time was about 11 minutes, at a rate of $3.50/hr.

Approximate charge to user = computer time + terminal time + communications + network overhead

= $0.17 + $0.64 + communications + network overhead

= $0.81 + communications + network overhead

CONTENTS—TEX-CODE

pages

1	Identification & Abstract
3	User Instructions
5– 6	I/O
7	Cost—Contents

DESCRIPTIVE TITLE	Fitting Nonlinear Regression Functions
CALLING NAME	TARSIER
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	J.D. Atkinson Statistical Laboratory Iowa State University
LANGUAGE	FORTRAN
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Available for use; not distribution
CONTACT	Louise R. Levine, Program Librarian Applications Programming, University Computer Center, The University of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

TARSIER is a programmed method for the fitting of nonlinear regression functions. The main program implements Hartley's[1] Modified Gauss—Newton Method for Fitting of Non-Linear Regression Functions by Least Squares. A subroutine, FCNT, is written for each individual application of TARSIER, describing the function to be fitted.

The Modified Gauss—Newton Method (MGN) is iterative and requires a starting value for each parameter to be estimated. The efficiency of the MGN method depends on the reasonable selection of the starting values.

REFERENCES

1. Hartley, H.O., "The Modified Gauss-Newton Method for the Fitting of Non-Linear Regression Functions by Least Squares," Technometrics, Vol. 3, No. 2 (May 1961), pp. 269—280. Reprint Series No. 90, Stat. Lab., Iowa State University.

continued

2. Hartley, H.O., and Booker, A., "Nonlinear Least Squares Estimation," Annals of Math. Stat., 36:2 (April 1965), pp. 638–650. Reprint Series No. 165, Stat. Lab., Iowa State University.

3. Faddev, D.K., *Computational Methods of Linear Algebra* (New York: Dover Publications Inc., 1959) pp. 105-111.

4. Atkinson, J.D., "TARSIER," (Ames, Iowa: Iowa State University Statistical Laboratory, 1969). Available from the EIN Office at the cost of reproduction and mailing.

USER INSTRUCTIONS

See the reference listed below.

REFERENCES

Atkinson, J.D., "TARSIER," (Ames, Iowa: Iowa State University Statistical Laboratory, 1969). Available from the EIN Office at the cost of reproduction and mailing.

SAMPLE INPUT

```
   1
HARTLEY MODIFIED GAUSS-NEWTON METHOD, ISU STAT LAB REPRINT NO. 90, PAGE 274
   4   1   3   6   2  10   0   0   2   1
(2F9.3)
127.     -5.
150.     -3.
379.     -1.
421.      1.
480.      3.
426.      5.
 31512.36
 32012.36
        .580E+03        -.180E+03        -.160E+00
        .500E+03        -.140E+03        -.180E+00
/*
```

SAMPLE OUTPUT

```
DATA SET NO.  1
                    HARTLEY MODIFIED GAUSS-NEWTON METHOD, ISU STAT LAB REPRINT NO. 90, PAGE 274



        ESTIMATE NO. 1 OF THE PARAMETERS


CYCLE               PARAMETER ESTIMATES, FOLLOWED BY SUM  OF SQUARES OF ERROR
   1     0.580000000000D 03     -0.180000000000D 03     -0.160000000000D 00      0.277090722845D 05
   2     0.508291052576D 03     -0.135599547036D 03     -0.215830015912D 00      0.152860528317D 05
   3     0.532100791582D 03     -0.162648053336D 03     -0.191893473223D 00      0.149357949293D 05
   4     0.535130406499D 03     -0.165401208689D 03     -0.195328448495D 00      0.147600883241D 05
   5     0.536193449611D 03     -0.166634580992D 03     -0.194217162209D 00      0.147597944574D 05



SOLUTION VECTOR IS
        THETA(I)

 0.53620409D 03

-0.16664071D 03

-0.19421612D 00



                X                          Y                      YHAT                        RESIDUAL
  -0.500000000000000D 01      0.127000000000000D 03      0.961398923965240 02      0.308601076034760 02
  -0.300000000000000D 01      0.150000000000000D 03      0.237788121398430 03     -0.877881213984290 02
  -0.100000000000000D 01      0.379000000000000D 03      0.333842501394140 03      0.451574986058630 02
   0.100000000000000D 01      0.421000000000000D 03      0.398978818179540 03      0.220211818204560 02
   0.300000000000000D 01      0.480000000000000D 03      0.443149003490810 03      0.368509965091940 02
   0.500000000000000D 01      0.426000000000000D 03      0.473101653052780 03     -0.471016530527830 02

RESIDUAL SUM OF SQUARES IS
RESID=  0.14759794D 05



        ESTIMATE NO. 2 OF THE PARAMETERS


CYCLE               PARAMETER ESTIMATES, FOLLOWED BY SUM  OF SQUARES OF ERROR
   1     0.500000000000D 03     -0.140000000000D 03     -0.180000000000D 00      0.205644162061D 05
   2     0.525982481594D 03     -0.156162828379D 03     -0.204548993581D 00      0.148010831208D 05
   3     0.535518330603D 03     -0.166377742162D 03     -0.194169172503D 00      0.147608563941D 05
   4     0.536183473170D 03     -0.166616755220D 03     -0.194237414849D 00      0.147597943973D 05
   5     0.536203750763D 03     -0.166640359513D 03     -0.194216430201D 00      0.147597942931D 05
```

continued

```
SCLUTION VECTOR IS
       THETA(I)

 0.536203790 03

-0.166640350 03

-0.194216440 00



          X                          Y                          YHAT                        RESIDUAL
  -0.500000000000000D 01    0.127000000000000D 03    0.961398293402120 02    0.308601706597880 02
  -0.300000000000000D 01    0.150000000000000D 03    0.237788172854560 03   -0.877881728545590 02
  -0.100000000000000D 01    0.379000000000000D 03    0.333842568489560 03    0.451574315104420 02
   0.100000000000000D 01    0.421000000000000D 03    0.398978853827550 03    0.220211461724470 02
   0.300000000000000D 01    0.480000000000000D 03    0.443148989296990 03    0.368510107030110 02
   0.500000000000000D 01    0.426000000000000D 03    0.473101585722540 03   -0.471015857225370 02

RESIDUAL SUM OF SQUARES IS
RESID=  0.147597940 05
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was approximately $1.61.

Charge to user = computer costs + postage + network overhead
= $1.61 + postage + network overhead

CONTENTS—TARSIER

pages
1– 2 Identification & Abstract
3 User Instructions
5– 7 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Student's t Test
CALLING NAME	STUDENTT
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Modified version version of BMDX70 (T Program) from U.C.L.A. Louise R. Levine University Computer Center The University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The University of Iowa, Iowa City, Iowa 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

This program computes t statistics and associated probability levels to test the equality of the means of two groups based on pooled and separate variance estimates. An F statistic and associated probability level for the equality of group variances is also computed. Paired comparison t ratios may be obtained through transgeneration. Groups are defined in two possible ways: 1) variable vs. variable, or 2) partitioned variable. Several dependent variables may be analyzed simultaneously. Each problem may contain from one to twenty subproblems. The cases to be included in each subproblem are determined through Boolean selection. Transgenerations are available and data specified as "missing" will be deleted. Variables transgenerated from variables with missing values will be considered missing.

Output

Output from this program includes F ratio of variances, t value (based on pooled variance estimate), t value (based on separate variance estimate), two-tailed probability levels for each t

continued

and for the F, means, standard deviations, standard error of the means, and number of observations included in the computation of the means, standard deviations, and standard error of the means.

Limitations per problem

p – number of original variables ($1 \le p \le 199$)

n – number of cases ($1 \le n \le 32{,}000$)

q – number of variables added to the original set after transgeneration ($-198 \le q \le 99$)

p+q – total number of variables output ($1 \le p+q \le 100$)

m – number of Transgeneration Cards ($0 \le m \le 100$)

D – number of Missing Value Cards ($0 \le D \le 100$)

b – number of Sub-problem Selection Cards per subproblem ($1 \le b \le 2$)

K – number of Variable Format Cards ($1 \le K \le 10$)

Machine Requirements

STUDENTT requires 124K bytes during execution and uses a special OS Assembler Language program, BLNK, for missing data.

USER INSTRUCTIONS

Input

The order of cards in the job deck should be as follows.

Problem Card
Missing Value Card(s) (optional)
Standard Transgeneration Card(s) (optional)
Subproblem Selection Card(s)
Category Selection Card(s)
F-type Variable Format Card(s)
Data Input Cards (Place data input deck here if data input is from cards)
Repeat for additional problems
Finish Card

Problem Card

Columns	*Contents*
1– 6	PROBLM
7–12	Alphanumeric problem code
13–17	Number of cases ($1 \le n \le 32,000$)
18–20	Number of original variables ($1 \le p \le 199$)
21–23	Number of Transgeneration Cards ($0 \le m \le 100$)
24–27	Number of variables added to or subtracted from the original set after transgeneration ($-198 \le q \le 99$)
28–30	Number of Missing Value Cards ($0 \le D \le 100$)
31–33	Number of subproblems This option must be used if the data are to be divided into two subgroups. This option must be included for variable vs. variable analysis, too. Each subproblem will compare the variable against *all* other variables. Only one Subproblem Card is needed, in either the variable vs. variable analysis or in the divided variable analysis.
34–36	Tape unit number to write data out for further processing. (The unit number must not be 5 or 6.) These columns may be blank if no more analysis is to be done.
37–38	Number of cards per case By specifying a tape number and the number of cards per case in Col. 34–36 and 37–38, respectively,

continued

471

Columns	*Contents*
	the user may write the input data (before transgeneration) on a specified logical tape for the processing of successive problems on the same input data. If a logical tape number is specified in Col. 34–36 of the Problem Card, the Problem Card of the following problem must have the same number in Col. 69–70. If this option is unnecessary, these columns should be left blank.
39–41	YES, if probabilities are to be computed. Otherwise leave blank.
42–44	YES, if PROBLM is variable-variable comparison. Otherwise leave blank.
69–70	00: data input is from cards T: data input is from logical tape $T(T \neq 6)$
71–72	Number of Variable Format Cards $(1 \leq K \leq 10)$
73–74	If Col. 42–44 = YES, the number of comparisons to be made (<20). Otherwise leave blank.

Missing Value Card(s) (optional)

If missing data on the card or tape are indicated by blanks or -0, they are detected without using these Missing Value Cards.

Columns	*Contents*
1– 4	Variable index, if card is for a single variable. If missing value codes are the same for all variables, punch 00 and prepare only one Missing Value Card.
5– 8	Number of missing value codes $(1 \leq C \leq 10)$
9–14	First missing value code
15–20	Second missing value code
⋮	⋮
63–68	Tenth missing value code

Standard Transgeneration Card(s) (optional)

A list of transgeneration codes is provided below. Variables transgenerated from variables with missing values will be considered missing. When a violation of a restriction in the

continued

right-hand column of the Transgeneration List occurs during transgeneration, the program will print a diagnostic message. After transgeneration of a particular case, if the original and added variables are not the first variables in the matrix (p+q) required for computation, they must be relocated. This may be done by using Transgeneration Code 09 with the constant c equal to one (Col. 7–9 on the TRNGEN Card should contain the new location and Col. 12–14 the previous location).

Columns	*Contents*
1– 6	TRNGEN
7– 9	Variable index k
10–11	Code from the Transgeneration List
12–14	Variable index i
15–20	Variable index j or constant c
21–25	Blank
26	Number of a_i's for transformation 40
27–32	a_1 value
33–38	a_2 value
⋮	⋮
63–68	a_7 value

The constants c, $a_1,\ldots,a_7$ are punched with a decimal point.

Example

To replace X_1 by exp $(-1/2X_1^2)$ three transformations are required.

```
TRNGEN001100012.0000
TRNGEN00109001-0.500
TRNGEN00104001000000
```

Transgeneration List

Notation to be used in the following transgeneration list:

i, j, k are variable indices (need not be different)

c is a constant

$a_1, a_2, a_3, \ldots$ are constants

continued

n is the number of cases, or sample size

the mean $\overline{X}_i = \frac{1}{n} \sum_{j=1}^{n} X_{ji}$

the standard deviation $s_i = \left[\frac{1}{n-1} \sum_{j=1}^{n} (X_{ji} - \overline{X}_i)^2\right]^{1/2}$

Code	*Transgeneration*	*Restriction*
01	$\sqrt{X_i} \rightarrow X_k$	$X_i \geq 0$
02	$\sqrt{X_i} + \sqrt{X_i + 1} \rightarrow X_k$	$X_i \geq 0$
03	$\log_{10} X_i \rightarrow X_k$	$X_i > 0$
04	$e^{X_i} \rightarrow X_k$	—
05	$\arcsin \sqrt{X_i} \rightarrow X_k$	$0 \leq X_i \leq 1$
07	$1/X_i \rightarrow X_k$	$X_i \neq 0$
08	$X_i + c \rightarrow X_k$	—
09	$X_i c \rightarrow X_k$	—
10	$X_i^c \rightarrow X_k$	$X_i \geq 0$
11	$X_i + X_j \rightarrow X_k$	—
12	$X_i - X_j \rightarrow X_k$	—
13	$X_i X_j \rightarrow X_k$	—
14	$X_i / X_j \rightarrow X_k$	$X_j \neq 0$
15	If $X_i \geq c$, $1 \rightarrow X_k$; otherwise $0 \rightarrow X_k$	—
16	If $X_i \geq X_j$, $1 \rightarrow X_k$; otherwise $0 \rightarrow X_k$	—
17	$\log_e X_i \rightarrow X_k$	$X_i > 0$

continued

Code	*Transgeneration*	*Restriction*
20	sin $X_i \rightarrow X_k$	—
21	cos $X_i \rightarrow X_k$	—
22	arctan $X_i \rightarrow X_k$	—
23	${X_i}^{X_j} \rightarrow X_k$	$X_i > 0$
24	$c^{X_i} \rightarrow X_k$	$c > 0$
40	If $X_i = a_1$ or a_2 or a_3 ..., a_7, then $c \rightarrow X_k$; otherwise X_k remains unchanged.	—
41	If X_i is blank, then $c \rightarrow X_k$; otherwise X_k remains unchanged	$(X_i \neq -0)$*

*Note that in reading numeric fields, a blank field and -0 are equivalent.

Subproblem Selection Card(s)

This card(s) indicates which cases are to be included in the subproblem.

Columns	*Contents*
1– 3	(V(
4– 6	Variable index for the first relationship
7	)
8– 9	Relationship: GT (Greater than) LT (Less than) GE (Greater than or equal to) LE (Less than or equal to) EQ (Equal to) NE (Not equal to)

continued

Columns	*Contents*
10–11	V(
12–14	Variable index of variable to be related to the preceding variable
15	)
16	blank
or	
10–15	Constant to be related to the preceding variable
16	)
17–18	Operation
	AN (and): The following relationship must also be true in order for the case to be included in this subproblem.
	OR: The case will be included if either the preceding or a following relationship is true.
	**: This terminates the set of Boolean relationships for this subproblem.

These formats (Col. 1–16) may be repeated up to four times per card ending in Col. 72. The maximum number of cards per subproblem is two; therefore, the user may specify up to eight relationships per subproblem. If four or more relationships are specified for a single subproblem, the first card must be complete. The first relationship for each subproblem must begin on a new card and the last operation must be **.

According to rule, the entire Boolean expression is either true or false for the case being tested. It is examined from left to right. If an OR is encountered and the expression preceding the OR is true, the case will be included in the subproblem. If the expression preceding the OR is false, the program begins again with the expression following the OR.

Examples

i. (V(002)NEV(100))**
The case is accepted if variable 2 is not equal to variable 100.

continued

ii. `(V(010)GE100.00)AN(V(010)LT200.0)**`
The case is accepted if variable 10 is greater than or equal to 100.00 and variable 10 is less than 200.00.

Category Selection Card(s)

Columns	*Contents*	
	Variable vs Variable	Partioned Variable
1– 3	first table variable no. one	first comparison variable no. first subproblem
4– 9	first table variable no. two	partition point for first subproblem (number punched here is the last included in first group)
10–12	second table variable no. one	All variables will be compared against each other. Only Col. 1–9 must be filled.
13–18	second table variable no. two	
⋮	⋮	
64–66	eighth table variable no. one	
67–72	eighth table variable no. two	

Continue in the same manner to name all comparisons. The total number of comparisons may not be greater than 20. The observations for a case not specified as missing are included in the X category if the value of the variable specified on the Category Selection Card is greater than or equal to the specified value; if the value is less, the case will be included in the Y category. If the category variable is missing, the case will be excluded from subproblems using that categorization.

Format Cards

These cards should be arranged by subject for all variables.

Finish Card (only one is used; next to last card)

Columns	*Contents*
1– 6	FINISH

A /* card is the last card in the deck.

SAMPLE INPUT

```
PRCBLM           4 18 18  18  C  1      YESYES                      00 116
TRNGEN 1911  1     4
TRNGEN 2211  2     5
TRNGEN 2511  3     6
TRNGEN 2011  7    10
TRNGEN 2311  8    11
TRNGEN 2611  9    12
TRNGEN 2111 13    16
TRNGEN 2411 14    17
TRNGEN 2711 15    18
TRNGEN 2814  2     1
TRNGEN 2914  5     4
TRNGEN 3014  8     7
TRNGEN 3114 11    10
TRNGEN 3214 14    13
TRNGEN 3314 17    16
TRNGEN 3414 22    19
TRNGEN 3514 23    20
TRNGEN 3614 24    21
(V(  1)LT     50)**
 25     27 20     21 23     24 26     27 28     29 30     31 32     33 29     30
 28     32 30     32 29     31 29     33 31     33 34     35 31     36 35     36
(16F5.0/2F5.0)
 8.8947,0655.9510.5832,4443,0212,0611.3123,37 9.4611,4620,92 3,6614-4218.08 6,14
 9.8716.01
 5.8722,4226.2912,46 7,9020.36 6,1413.9520,09 4.4020.4124,8110,7820,5931.3714,41
34.2148,67
14.2927,9542-22 6.3939,2245.6411,6316.6428,2712.7214.2626,98

 7.6626,8034,46

FINISH
/*
```

478

SAMPLE OUTPUT

```
BMDX70 T PROGRAM - REVISED OCTOBER 24, 1968
HEALTH SCIENCES COMPUTING FACILITY,UCLA

PROBLEM
NUMBER OF CASES..........     4   NO. OF SUB-PROBLEMS. 1
NUMBER OF PUNCHED VARIABLES. 18   TAPE TO WRITE....... 0
NUMBER OF TRANSGENERATIONS.. 18   INPUT TAPE NUMBER... 0
NUMBER OF VARIABLES ADDED... 18   NUMBER OF VARIABLE
NUMBER OF MISVAL CARDS......  0   FORMAT CARDS........ 1

 NEW            OLD     B VAR.    NUMBER OF
 VAR.   CODE    VAR.   OR CONST.  CONSTANTS        TYPE 40 CONSTANTS
  19     11      1        4.00
  22     11      2        5.00
  25     11      3        6.00
  20     11      7       10.00
  23     11      8       11.00
  26     11      9       12.00
  21     11     13       16.00
  24     11     14       17.00
  27     11     15       18.00
  28     14      2        1.00
  29     14      5        4.00
  30     14      8        7.00
  31     14     11       10.00
  32     14     14       13.00
  33     14     17       16.00
  34     14     22       19.00
  35     14     23       20.00
  36     14     24       21.00

    VARIABLE FORMAT
(16F5.0/2F5.0)



SUB-PROBLEM NUMBER  1 CONTAINS     4 CASES                         *  POOLED VARIANCE ESTIMATE   *  SEPARATE VARIANCE ESTIMATE
                                                         *         *                             *
VARIABLE    NUMBER                STANDARD    STANDARD   *   F      P     *    T       DEGREES     P     *    T      DEGREES       P
 INDEX     OF CASES      MEAN     DEVIATION     ERROR    * VALUE  VALUE   *  VALUE   OF FREEDOM  VALUE   *  VALUE  OF FREEDOM    VALUE
**********************************************************************************************************************************
                                                         *                *                              *
                                                         *                *                              *
  25          4        58.4327     44.845      22.422    *                *                              *
                                                         *  1.91  0.957   *   0.04         4     0.971   *   0.04       2.93     0.969
  27          2        57.0400     32.456      22.950    *                *                              *
                                                         *                *                              *
----------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER  2 CONTAINS     4 CASES                         *  POOLED VARIANCE ESTIMATE   *  SEPARATE VARIANCE ESTIMATE
                                                         *         *                             *
VARIABLE    NUMBER                STANDARD    STANDARD   *   F      P     *    T       DEGREES     P     *    T      DEGREES       P
 INDEX     OF CASES      MEAN     DEVIATION     ERROR    * VALUE  VALUE   *  VALUE   OF FREEDOM  VALUE   *  VALUE  OF FREEDOM    VALUE
**********************************************************************************************************************************
                                                         *                *                              *
                                                         *                *                              *
  20          3        18.8033      7.295       4.212    *                *                              *
                                                         *  2.23  0.857   *   0.17         3     0.879   *   0.15       1.62     0.906
  21          2        17.4950     10.882       7.695    *                *                              *
                                                         *                *                              *
----------------------------------------------------------------------------------------------------------------------------------
```

000 0129

continued

```
SUB-PROBLEM NUMBER  3 CONTAINS     4 CASES                        *  POOLED VARIANCE ESTIMATE     *  SEPARATE VARIANCE ESTIMATE
                                                                  *                               *
VARIABLE   NUMBER              STANDARD    STANDARD  *   F      P    *   T       DEGREES      P    *   T       DEGREES      P
 INDEX    OF CASES     MEAN    DEVIATION     ERROR   *  VALUE  VALUE *  VALUE   OF FREEDOM  VALUE  *  VALUE   OF FREEDOM   VALUE
*************************************************************************************************************************************
                                                     *               *                             *
   23        3      29.3433      5.950       3.435   *               *                             *
                                                     *  13.15  0.383 *  -0.84        3     0.465   *  -0.65       1.10      0.632
   24        2      39.5450     21.574      15.255   *               *                             *
                                                     *               *                             *
-------------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER  4 CONTAINS     4 CASES                        *  POOLED VARIANCE ESTIMATE     *  SEPARATE VARIANCE ESTIMATE
                                                                  *                               *
VARIABLE   NUMBER              STANDARD    STANDARD  *   F      P    *   T       DEGREES      P    *   T       DEGREES      P
 INDEX    OF CASES     MEAN    DEVIATION     ERROR   *  VALUE  VALUE *  VALUE   OF FREEDOM  VALUE  *  VALUE   OF FREEDOM   VALUE
*************************************************************************************************************************************
                                                     *               *                             *
   26        3      48.1467      6.159       3.556   *               *                             *
                                                     *  27.77  0.266 *  -0.50        3     0.650   *  -0.38       1.05      0.767
   27        2      57.0400     32.456      22.950   *               *                             *
                                                     *               *                             *
-------------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER  5 CONTAINS     4 CASES                        *  POOLED VARIANCE ESTIMATE     *  SEPARATE VARIANCE ESTIMATE
                                                                  *                               *
VARIABLE   NUMBER              STANDARD    STANDARD  *   F      P    *   T       DEGREES      P    *   T       DEGREES      P
 INDEX    OF CASES     MEAN    DEVIATION     ERROR   *  VALUE  VALUE *  VALUE   OF FREEDOM  VALUE  *  VALUE   OF FREEDOM   VALUE
*************************************************************************************************************************************
                                                     *               *                             *
   28        4       3.5567      1.365       0.682   *               *                             *
                                                     *   4.08  0.406 *   0.18        5     0.866   *   0.16       2.74      0.887
   29        3       3.2793      2.758       1.592   *               *                             *
                                                     *               *                             *
-------------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER  6 CONTAINS     4 CASES                        *  POOLED VARIANCE ESTIMATE     *  SEPARATE VARIANCE ESTIMATE
                                                                  *                               *
VARIABLE   NUMBER              STANDARD    STANDARD  *   F      P    *   T       DEGREES      P    *   T       DEGREES      P
 INDEX    OF CASES     MEAN    DEVIATION     ERROR   *  VALUE  VALUE *  VALUE   OF FREEDOM  VALUE  *  VALUE   OF FREEDOM   VALUE
*************************************************************************************************************************************
                                                     *               *                             *
   30        3       1.5469      0.675       0.389   *               *                             *
                                                     *   8.84  0.293 *  -0.64        4     0.559   *  -0.64       2.45      0.59
   31        3       2.3237      2.005       1.158   *               *                             *
                                                     *               *                             *
-------------------------------------------------------------------------------------------------------------------------------------
```

continued

```
SUB-PROBLEM NUMBER  7 CONTAINS     4 CASES                          *  POOLED VARIANCE ESTIMATE   *  SEPARATE VARIANCE ESTIMATE
                                                          *                   *                             *
VARIABLE   NUMBER               STANDARD    STANDARD    *    F       P    *    T       DEGREES      P    *    T       DEGREES       P
 INDEX   OF CASES      MEAN    DEVIATION      ERROR     *  VALUE   VALUE  *  VALUE   OF FREEDOM   VALUE  *  VALUE   OF FREEDOM    VALUE
*********************************************************************************************************************************
                                                          *                   *                             *
  32         2       2.9250      1.435        1.015     *                   *                             *
                                                          *  7.01   0.460   *  0.86          2    0.480   *  0.86         1.28     0.547
  33         2       1.9908      0.542        0.383     *                   *                             *
                                                          *                   *                             *
---------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER  8 CONTAINS     4 CASES                          *  POOLED VARIANCE ESTIMATE   *  SEPARATE VARIANCE ESTIMATE
                                                          *                   *                             *
VARIABLE   NUMBER               STANDARD    STANDARD    *    F       P    *    T       DEGREES      P    *    T       DEGREES       P
 INDEX   OF CASES      MEAN    DEVIATION      ERROR     *  VALUE   VALUE  *  VALUE   OF FREEDOM   VALUE  *  VALUE   OF FREEDOM    VALUE
*********************************************************************************************************************************
                                                          *                   *                             *
  28         4       3.5567      1.365        0.682     *                   *                             *
                                                          *  4.09   0.405   *  2.31          5    0.069   *  2.56         4.55     0.063
  30         3       1.5469      0.675        0.389     *                   *                             *
                                                          *                   *                             *
---------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER  9 CONTAINS     4 CASES                          *  POOLED VARIANCE ESTIMATE   *  SEPARATE VARIANCE ESTIMATE
                                                          *                   *                             *
VARIABLE   NUMBER               STANDARD    STANDARD    *    F       P    *    T       DEGREES      P    *    T       DEGREES       P
 INDEX   OF CASES      MEAN    DEVIATION      ERROR     *  VALUE   VALUE  *  VALUE   OF FREEDOM   VALUE  *  VALUE   OF FREEDOM    VALUE
*********************************************************************************************************************************
                                                          *                   *                             *
  28         4       3.5567      1.365        0.682     *                   *                             *
                                                          *  1.11   1.176   *  0.53          4    0.626   *  0.52         1.97     0.696
  32         2       2.9250      1.435        1.015     *                   *                             *
                                                          *                   *                             *
---------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER 10 CONTAINS     4 CASES                          *  POOLED VARIANCE ESTIMATE   *  SEPARATE VARIANCE ESTIMATE
                                                          *                   *                             *
VARIABLE   NUMBER               STANDARD    STANDARD    *    F       P    *    T       DEGREES      P    *    T       DEGREES       P
 INDEX   OF CASES      MEAN    DEVIATION      ERROR     *  VALUE   VALUE  *  VALUE   OF FREEDOM   VALUE  *  VALUE   OF FREEDOM    VALUE
*********************************************************************************************************************************
                                                          *                   *                             *
  30         3       1.5469      0.675        0.389     *                   *                             *
                                                          *  4.53   0.631   * -1.52          3    0.227   * -1.27         1.30     0.425
  32         2       2.9250      1.435        1.015     *                   *                             *
                                                          *                   *                             *
---------------------------------------------------------------------------------------------------------------------------------
```

continued

000 0129

```
SUB-PROBLEM NUMBER 11 CONTAINS      4 CASES                              * POOLED VARIANCE ESTIMATE        * SEPARATE VARIANCE ESTIMATE
                                                             *                  *                                 *
VARIABLE     NUMBER                 STANDARD     STANDARD    *    F       P     *    T       DEGREES        P     *    T      DEGREES         P
 INDEX     OF CASES       MEAN     DEVIATION      ERROR      *  VALUE   VALUE   *  VALUE   OF FREEDOM    VALUE  *  VALUE   OF FREEDOM     VALUE
*****************************************************************************************************************************************************
                                                             *                  *                                 *
                                                             *                  *                                 *
  29          3         3.2793      2.758        1.592       *                  *                                 *
                                                             *   1.89   0.692   *   0.49         4        0.653   *   0.49       3.65       0.661
  31          3         2.3237      2.005        1.158       *                  *                                 *
                                                             *                  *                                 *
-----------------------------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER 12 CONTAINS      4 CASES                              * POOLED VARIANCE ESTIMATE        * SEPARATE VARIANCE ESTIMATE
                                                             *                  *                                 *
VARIABLE     NUMBER                 STANDARD     STANDARD    *    F       P     *    T       DEGREES        P     *    T      DEGREES         P
 INDEX     OF CASES       MEAN     DEVIATION      ERROR      *  VALUE   VALUE   *  VALUE   OF FREEDOM    VALUE  *  VALUE   OF FREEDOM     VALUE
*****************************************************************************************************************************************************
                                                             *                  *                                 *
                                                             *                  *                                 *
  29          3         3.2793      2.758        1.592       *                  *                                 *
                                                             *  25.89   0.275   *   0.62         3        0.579   *   0.79       2.22       0.514
  33          2         1.9908      0.542        0.383       *                  *                                 *
                                                             *                  *                                 *
-----------------------------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER 13 CONTAINS      4 CASES                              * POOLED VARIANCE ESTIMATE        * SEPARATE VARIANCE ESTIMATE
                                                             *                  *                                 *
VARIABLE     NUMBER                 STANDARD     STANDARD    *    F       P     *    T       DEGREES        P     *    T      DEGREES         P
 INDEX     OF CASES       MEAN     DEVIATION      ERROR      *  VALUE   VALUE   *  VALUE   OF FREEDOM    VALUE  *  VALUE   OF FREEDOM     VALUE
*****************************************************************************************************************************************************
                                                             *                  *                                 *
                                                             *                  *                                 *
  31          3         2.3237      2.005        1.158       *                  *                                 *
                                                             *  13.69   0.375   *   0.22         3        0.841   *   0.27       2.40       0.810
  33          2         1.9908      0.542        0.383       *                  *                                 *
                                                             *                  *                                 *
-----------------------------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER 14 CONTAINS      4 CASES                              * POOLED VARIANCE ESTIMATE        * SEPARATE VARIANCE ESTIMATE
                                                             *                  *                                 *
VARIABLE     NUMBER                 STANDARD     STANDARD    *    F       P     *    T       DEGREES        P     *    T      DEGREES         P
 INDEX     OF CASES       MEAN     DEVIATION      ERROR      *  VALUE   VALUE   *  VALUE   OF FREEDOM    VALUE  *  VALUE   OF FREEDOM     VALUE
*****************************************************************************************************************************************************
                                                             *                  *                                 *
                                                             *                  *                                 *
  34          4         2.4691      1.440        0.720       *                  *                                 *
                                                             *   1.41   0.883   *   0.59         5        0.583   *   0.60       4.85       0.579
  35          3         1.8623      1.215        0.701       *                  *                                 *
                                                             *                  *                                 *
-----------------------------------------------------------------------------------------------------------------------------------------------------
```

continued

```
SUB-PROBLEM NUMBER 15 CONTAINS     4 CASES                               * POOLED VARIANCE ESTIMATE      * SEPARATE VARIANCE ESTIMATE
                                                            *                 *                                *
VARIABLE   NUMBER              STANDARD    STANDARD   *    F       P      *    T      DEGREES       P     *    T      DEGREES       P
  INDEX   OF CASES      MEAN   DEVIATION     ERROR    *  VALUE   VALUE    *  VALUE   OF FREEDOM   VALUE   *  VALUE   OF FREEDOM   VALUE
*************************************************************************************************************************************
                                                      *                   *                               *
                                                      *                   *                               *
   31         3       2.3237     2.005       1.158    *                   *                               *
                                                      *  87.53  0.151     *  -0.00          3      0.998  *  -0.00        2.07     0.998
   36         2       2.3270     0.214       0.152    *                   *                               *
                                                      *                   *                               *
-------------------------------------------------------------------------------------------------------------------------------------


SUB-PROBLEM NUMBER 16 CONTAINS     4 CASES                               * POOLED VARIANCE ESTIMATE      * SEPARATE VARIANCE ESTIMATE
                                                            *                 *                                *
VARIABLE   NUMBER              STANDARD    STANDARD   *    F       P      *    T      DEGREES       P     *    T      DEGREES       P
  INDEX   OF CASES      MEAN   DEVIATION     ERROR    *  VALUE   VALUE    *  VALUE   OF FREEDOM   VALUE   *  VALUE   OF FREEDOM   VALUE
*************************************************************************************************************************************
                                                      *                   *                               *
                                                      *                   *                               *
   35         3       1.8623     1.215       0.701    *                   *                               *
                                                      *  32.13  0.248     *  -0.51          3      0.646  *  -0.65        2.18     0.584
   36         2       2.3270     0.214       0.152    *                   *                               *
                                                      *                   *                               *
-------------------------------------------------------------------------------------------------------------------------------------


NUMBER OF CASES NOT INCLUDED IN DESIGNATED SUB-PROBLEMS=     0
NUMBER OF INSTANCES IN WHICH TRANSGENERATION RESTRICTIONS WERE VIOLATED =     0
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, I/O accesses, core storage, cards in and out, and lines and pages printed. The total cost for the Sample Output above was $3.82.

Charge to user = computer costs + postage + network overhead
= $3.82 + postage + network overhead

CONTENTS—STUDENTT

pages
1– 2 Identification & Abstract
3– 9 User Instructions
11–16 I/O
17 Cost—Contents

DESCRIPTIVE TITLE	Correlation and Multiple Linear Regression
CALLING NAME	UMST500
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program performs a correlation analysis followed by a multiple linear regression. This program does *not* handle missing data. Any or all of the following may be calculated and given as output:

1. Raw cross-product matrix
2. Cross-product matrix about the means
3. Covariance matrix
4. Means and standard deviations
5. Correlation coefficient matrix
6. Multiple correlation coefficient
7. Ordinary regression coefficients and their standard errors
8. Normal regression coefficients and their standard errors
9. Total sum of squares of the dependent variable about its mean, separated into regression sum of squares and error sum of squares
10. Standard error of estimate
11. t statistics for testing significance of regression coefficients (regression coefficient divided by the standard error of that coefficient)
12. Partial correlation coefficients of the dependent variable with the independent variables

continued

13. Inverse of correlation matrix with dependent variable
14. Check on accuracy of inverse
15. Covariance matrix of normal regression coefficients
16. Inverse of cross-product matrix (about means) without the dependent variable
17. Back solution: observed dependent variable minus predicted dependent variable (residual) for all observations (The Durbin-Watson statistic is output when this option is chosen.)

Seven different transformations of all variables can be made. Twenty-one methods of transformation and/or generation of individual variables are possible. (See User Instructions: Problem Card and Transgeneration Cards.)

Restrictions on J, the number of variables:

$J \leq 216$ if only a correlation solution is called for; i.e., outputs 1–5 above.
$J \leq 145$ if a regression solution is called for, excluding outputs 14–16 above.
$J \leq 90$ if a regression solution is called for including all the above; i.e., outputs 1–17.

Restrictions on N, the number of observations (or individuals)

$N \leq 2^{16}-1 = 32{,}767$ if N is supplied to the program by an input card.
$N \leq 2^{47}-1 \doteq 10^{14}$ if the program is to count the number of observations as they are read in.

This program can compute any multiple regression equation which involves variables contained in the correlation matrix (output 5 above); 6–17 may be computed for each such equation. Any variable may be named as the dependent variable.

An optional feature of this program is the calculation of all possible regression equations involving *one dependent variable and up to 10 independent variables*. For example, if y is the dependent variable and x_1, x_2, x_3, are independent variables, then the following regressions would be obtained:

y and x_1	y and x_1, x_2
y and x_2	y and x_1, x_3
y and x_3	y and x_2, x_3

y and x_1, x_2, x_3

continued

An additional option is the use of a special calculation technique to compute the cross-product matrix (about means), without possible loss of significant digits, directly from the input data.

REFERENCES

Anderson, D. and Frisch, M., *Statistical Programs for Use on the CDC 6600 Computer* (Minneapolis, Minn.: Univ. of Minn., Univ. Comp. Ctr., 1969).

Welford, B.P., "Note on a Method for Calculating Corrected Sums of Squares and Products", Technometrics, *4* (1962), pp. 419-420.

Caffrey, J., "ALGORITHM 66", Communications of the Association for Computing Machinery, *4* (1961), p. 322.

Johnson, P. O., *Statistical Methods in Research* (New York: Prentice Hall, Inc., 1949), pp. 338-340.

A Multiple Regression and Correlation Program for the Univac Scientific (1103), RR–127 (New York: UNIVAC, 1956).

USER INSTRUCTIONS

System Control Cards (will be prepared by Univ. of Minn. personnel.)

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8–10	NVAR	Number of variables to be read from the Data Cards (does not include optional counting/card-ordering variable). Check this number against restrictions on J, above. NVAR $\geq$ 2
11–15	NOBS	0: computer counts observations without checking card ordering. 1: computer counts observations and checks on card ordering (see Input Data Cards). NOBS: exactly the number of observations the computer is to read in. NOBS $\geq$ 0
16	NFC	Number of Format Cards ($\leq$ 5) If zero or blank, NFC = 1 is assumed.
17		1: back solution is required in a regression problem performed on this data (this causes data to be saved on magnetic tape). 0: otherwise
18		1: use special method for direct calculation of cross-product matrix (about means), omitting raw cross-product matrix 0: use standard method for calculation of raw cross-product matrix
19		1: output correlation coefficients 0: otherwise
20		1: output standard deviations and means 0: otherwise
21		1: output covariance matrix 0: otherwise

continued

Columns	*Parameter*	*Contents*
22		1: output of cross-product matrix (about means) 0: otherwise
23		1: output of raw cross-product matrix 0: otherwise
24		0: 4 place accuracy 1: 7 place accuracy 2: ten place accuracy (in the printout of cross-product and covariance matrices)
25–27	NRG	Number of regression equations to be computed from this data (i.e., the number of REGRESS Cards)
28–30	NVG	Number of Transgeneration Cards ($\leq$ 250)
31	NTRANS	0: no transformations are to be made 1: $\log_{10}$ transformation of all variables 2: $\log_e$ transformation of all variables 3: square root transformation of all variables 4: square transformation of all variables 5: reciprocal transformation of all variables 6: sine transformation of all variables 7: cosine transformation of all variables
32–34	NADD	The number of new variables added by transgeneration. See Transgeneration Cards below.
35–80	Name	Alphanumeric identification of the problem

Format Card(s)

Variables are considered to be indexed, or numbered, by a subscript according to the order in which they are read in from the Data Cards. All data must be read with F, E, or X – fields.

continued

The first variable read is No. 1, the second is No. 2, etc., and they are denoted by x_1, x_2, ..., x_j. The Format Card must provide for reading in the NVAR elements of an observation at one time. In addition, if either 0 or 1 is punched in Col. 11–15 of the Problem Card, the first field read by the Format Card shall be an integer field of at least 2 columns. This field is used in checking card order and counting observations.

Transgeneration Cards

Each Transgeneration Card specifies the transformation or variable generation desired. The new variable or transformed variable may replace one of the existing variables or may be added to the original set of variables. Transgenerations are performed sequentially; thus, one can add two variables in one transgeneration and take the square root of the sum in the next transgeneration. Where there are NVAR variables originally introduced into the problem, we shall indicate by NADD the total additional variables occupying new variable locations which are to be used subsequently. Do not include in NADD the transgenerations which replace other variables. However, the limit on NVAR and NADD specified by the program cannot be exceeded even at intermediate stages of the transgeneration. Transformations of all variables at once as indicated by Col. 31 of the Problem Card operate independently of and after transgeneration and may be used alternatively or in conjunction with Transgeneration Cards. Violation of restrictions relating to the transformations and transgenerations results in omission of all variables of the observation for which the violation occurred, and in an appropriate print-out indicating the observation omitted. *Note:* The sum of Col. 8 – 10 and Col. 32 – 34 on the Problem Card must satisfy the restrictions on J.

Notation

x:	the variable prior to transgeneration
x':	the transformed variable which (depending on instructions in the transgenerator card) replaces the original variable x or becomes a new variable
c:	a constant
A, B:	specified variable numbers, or indices
*:	multiplication

Columns	*Contents*
1– 3	Variable number to be assigned to x'
4– 5	Transgenerator code

continued

Columns	*Contents*
6–10	A: variable number, or variable number of x, or constant C_1 for transformation 21
11–20	B: variable number, or constant c, or constant C_2 for transformation 21 (keypunch decimal point if not at extreme right)

Code	*Transformation*	*Restriction*
00	$x' = x$	
01	$x' = \sqrt{x}$	$x \geq 0$
02	$x' = \sqrt{x} + \sqrt{x + 1}$	$x \geq 0$
03	$x' = \log_{10} x$	$x > 0$
04	$x' = e^x$	$x \leq 740$
05	$x' = \sin^{-1} x$	$-1 \leq x \leq 1$
06	$x' = \log_e x$	$x > 0$
07	$x' = 1/x$	$x \neq 0$
08	$x' = x + c$	
09	$x' = x * c$	
10	$x' = x^c$	$x \geq 0$
11	$x' = x_A + x_B$	
12	$x' = x_A - x_B$	
13	$x' = x_A * x_B$	
14	$x' = x_A / x_B$	$x_B \neq 0$
15	if $x \geq c$, set $x' = 1$ otherwise, set $x' = 0$	
16	if $x_A \geq x_B$, set $x' = 1$ otherwise, set $x' = 0$	
17	$x' = \sin x$	$\lvert x \rvert < 2.2 \times 10^{14}$ x in radians
18	$x' = \cos x$	$\lvert x \rvert < 2.2 \times 10^{14}$ x in radians
19	$x' = {x_A}^{x_B}$	$x_A > 0$
20	$x' = c^x$	$c > 0$
21	if $x = C_1$, set $x' = 1$ if $x = C_2$, set $x' = -1$ otherwise, set $x' = 0$	

continued

Code	*Transformation*	*Restriction*
Note on Code 21:	[Notice that the x variable (no. indicated in Col. 1–3) is replaced by 1, –1, or 0. If you do not want to destroy your original variable, it is necessary to create a new one by using transformation 00, and then using transformation 21 on the new variable.]	

Input Data Cards

The Data Cards shall be punched in accordance with the format given on the Format Card. The NVAR elements of an observation are read at one time from a set of cards. Each observation starts on a new set of cards. If 0 is punched in Col. 11 – 15 of the Problem Card, there must be a non-negative integer in the first field (an I-field) read by the Format Card, for use in counting observations. If 1 is punched in Col. 11 – 15 of the Problem Card, there must be a 1 in the first field (an I-field) read from each observation by the Format Card, for use in checking card ordering. This assumes more than one Data Card per observation and each is numbered 1, 2, 3, etc. Thus, if a 2 or 3 is found, there would be an error in card ordering.

End of Data Card

This card is required only if Col. 11 – 15 of the Problem Card contains 0 or 1. It must have –1 punched where it will be read by the first (i.e. integer) field of the Format Card. If more than one data card is read at one time, this card must be followed by enough blanks to make a set of Data Cards.

Regression Cards

REGRESS Cards

(There must be as many REGRESS Cards as indicated by Col. 25 – 27 of the Problem Card.)

Columns	*Contents*
1– 7	REGRESS
8–12	Numeric identification for problems based on this card

continued

Columns	*Contents*
13	0: No back solution is wanted 1: Print back solution for all equations based on this card 2: Print back solution for final equation only

The following three sets of two column options are mutually exclusive. That is, you may choose the options in 14 – 15 or 16 – 17 or 18 – 19, with Col. 16 – 17 being the most likely choice.

Columns	*Contents*
14	1: Print all possible regression solutions with x_1 as dependent variable, with output as indicated by Col. 15. If Col. 14 is punched 1, then this Regress Card must be the last one 0: otherwise
15	1: Print normal regression coefficients and standard errors 2: Print ordinary regression coefficients, standard errors, t's, and partial correlations of dependent variable with each of the independent variables 3: Print both of the above 0: Print only outputs 6, 9, 10 as indicated in Abstract, and indices of dependent and independent variables
16	1: List of dependent and independent variables follows on this card; output as indicated by Col. 17 0: otherwise
17	Same codes as Col. 15
18	1: Print output for the regression which includes all variables, in accordance with Col. 19. Variable No. 1 is the dependent variable in this case. If Col. 18 is punched 1, then this Card must be the last one. 0: otherwise
19	Same codes as Col. 15.

continued

Columns	*Contents*
20	See restrictions on J in Functional Abstract

	Print inverse of correlation matrix with dependent variable for:	Print inverse check for:
0:	no equations	no equations
1:	all equations	no equations
2:	final equation only	no equations
3:	no equations	all equations
4:	all equations	all equations
5:	final equation only	all equations
6:	no equations	final equation only
7:	all equations	final equation only
8:	final equation only	final equation only

Columns	*Contents*
21	See restrictions on J in abstract

Print covariance matrix of normal coefficients for:	Print inverse of cross-product matrix (about means) without dependent variable for:
Same codes as Col. 20	Same codes as Col. 20

Columns	*Contents*
22–24	Number of indices to be read from this list (including dependent variable) which follows starting in Col. 25
25–28	The index of the dependent variable for this regression (*Note:* the first variable on a Data Card is considered to be No. 1, the second No. 2, etc.)
29–32	The index of the first independent variable
33–36	The index of the second independent variable
⋮	⋮
77–80	The index of the 13th independent variable.

Note: A minus sign preceeding a variable index will cause the intermediate equation which first includes that variable to be printed; for example, the list starting in Col. 25 – 28 which reads 6 4 –2 1 3 will give 2 regressions with the 6th variable as dependent:

6 with 4 and 2
6 with 4,2,1, and 3

continued

500

494

CONTINU Card(s)

(these cards are required only as continuation for the index list on the REGRESS Card)

Columns	*Contents*
1– 7	CONTINU
8	blank
9–80	Continuation of the list of variable numbers used in this regression, in 4–column fields as before.

Finish Card

Columns	*Contents*
1– 6	FINISH

Used after last problem (see next section)

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

Program Control Cards for problem 1
    Problem Card for problem 1
    Format Card(s) for problem 1
    Transgeneration Cards (if any) for problem 1
Data Cards for problem 1
End of Data Card for problem 1 (if required)
Regression Cards for problem 1 (if required)
    REGRESS Card
    CONTINU Card(s) (if required)
Program Control Cards for problem 2
Data Cards for problem 2
End of Data Card for problem 2 (if required)
Regression Card(s) for problem 2 (if required)
  ⋮
Program Control Cards for last problem
Data Cards for last problem
End of Data Card for last problem (if required)
Regression Card(s) for last problem (if required)
Finish Card

Further discussion of error printouts, notation, and computation can be found in Ref. 1.

continued

REFERENCES

1. Anderson, D. and Frisch, M., *Statistical Programs for Use on the CDC 6600 Computer* (Minneapolis, Minn.: Univ. of Minn., Univ. Comp. Ctr., 1969).

SAMPLE INPUT

```
PROBLEM  3       110111111  1  10   1UMST500 SAMPLE PROBLEM
(I2,3F2.0)
  413    2          3
   3 6 5
   1 2 3
   2 3 4
   6 5 4
   3 4 5
   1 3 2
   2 4 3
   3 5 4
-1
REGRESS    12  13  48  4   2   3  -4   1
FINISH
```

SAMPLE OUTPUT

```
UMST500          CORRELATION AND REGRESSION ANALYSIS  UMST500 SAMPLE PROBLEM


NVAR =   3  NOBS =   -0  NFC = 1  NVG =  1   NADD =  1   NTRANS = 0   NRG =  1
    OPTIONS          CORRELATION COEFFICIENTS,   YES          MEANS AND ST DEV,   YES
                     COVARIANCE MATRIX,   YES                 CROSS-PRODUCT MATRIX(ABOUT MEANS),   YES
                     RAW CROSS-PRODUCT MATRIX,   YES          ACCURACY,    7 PLACE
                     BACK SOLUTION,   YES                     SPECIAL METHOD,    NO
        (MATRICES ARE PRINTED BY ROWS OF THE LOWER LEFT TRIANGLE, INCLUDING DIAGONAL ELEMENTS)


         FORMAT = (I2,3F2.0)

      TRANS-GENERATOR CARD(S)

CARD    NEW     TRANS    ORIG.   ORIG. VAR(B)
 NO. VARIABLE   CODE    VAR(A)   OR CONSTANT
 1       4       13      2.00          3.000000000

    NO. OF OBSERVATIONS =      8


       COVARIANCE MATRIX

1      2.553571E+00

2      1.428571E+00  1.714286E+00

3      8.928571E-01  8.571429E-01  1.071429E+00

4      8.464286E+00  1.000000E+01  7.357143E+00  6.792857E+01


CROSS-PRODUCT MATRIX (RAW),  COLUMN J IN LAST ROW IS SUM OF X(J)
   1      7.300000E+01

   2      9.400000E+01  1.400000E+02

   3      8.500000E+01  1.260000E+02  1.200000E+02

   4      3.900000E+02  5.740000E+02  5.240000E+02  2.460000E+03

   5      2.100000E+01  3.200000E+01  3.000000E+01  1.260000E+02


       CROSS-PRODUCT MATRIX(ABOUT MEANS)

1      1.787500E+01

2      1.000000E+01  1.200000E+01

3      6.250000E+00  6.000000E+00  7.500000E+00

4      5.925000E+01  7.000000E+01  5.150000E+01  4.755000E+02
```

continued

```
                MEAN            STANDARD DEVIATION
1        2.6250000000E+00     1.5979898087E+00
2        4.0000000000E+00     1.3093073414E+00
3        3.7500000000E+00     1.0350983390E+00
4        1.5750000000E+01     8.2418791200E+00


      CORRELATION COEFFICIENTS

1       1.000000

2        .682789  1.000000

3        .539792   .632456  1.000000

4        .642673   .926685   .862385  1.000000



.........REGRESSION EQUATION      1....................UMST500 SAMPLE PROBLEM                              ....

                           DEPENDENT VARIABLE   2                     MULTIPLE  CORR  COEFF
                             2 INDEP VARIABLES                        1-R2     R2      R
                             5 DEG OF FREEDOM                         .03282  .96718  .98345

          NORMAL EQUATION                         SUM OF SQUARES                               STANDARD ERROR OF ESTIMATE
                                        TOTAL         REGR           ERROR                       SE2            SE
                                     7.00000E+00   6.77024E+00    2.29759E-01                4.59519E-02    2.14364E-01

                 3            4
B          -6.5045E-01  1.4876E+00
SD          1.6004E-01  1.6004E-01

          ORDINARY EQUATION                       SUM OF SQUARES                               STANDARD ERROR OF ESTIMATE
                                        TOTAL         REGR           ERROR                       SE2            SE
                                     1.20000E+01   1.16061E+01    3.93873E-01                7.87746E-02    2.80668E-01
 CONSTANT        3.3632385120E+00

                 3            4
B          -8.2276E-01  2.3632E-01
SD          2.0244E-01  2.5424E-02
B/SD       -4.0642E+00  9.2952E+00
PCYI       -8.7615E-01  9.7226E-01

INVERSE OF CORRELATION MATRIX (WITH DEPENDENT VARIABLE). SIGN IS REVERSED

    1         -3.046666667E+01

    2         -1.981694000E+01  -1.679166667E+01

    3          4.532284192E+01   3.284494824E+01  -7.132500000E+01

INVERSE CHECK

    1          1.000000000E+00

    2         -2.273736754E-13   1.000000000E+00

    3         -4.547473509E-13  -2.273736754E-13   1.000000000E+00
```

continued

```
.........REGRESSION EQUATION     1...................UMST500 SAMPLE PROBLEM                                .....

                     DEPENDENT VARIABLE   2                       MULTIPLE  CORR  COEFF
                       3 INDEP VARIABLES                          1-R2     R2      R
                       4 DEG OF FREEDOM                           .02249  .97751  .98869

          NORMAL EQUATION                          SUM OF SQUARES                             STANDARD ERROR OF ESTIMATE
                                        TOTAL          REGR          ERROR                     SE2            SF
                                     7.00000E+00   6.84260E+00   1.57399E-01                3.93498E-02   1.98368E-01

                 3            4            1
B          -6.4296E-01  1.3958E+00  1.3280E-01
SD          1.4820E-01  1.6284E-01  9.7930E-02

          ORDINARY EQUATION                        SUM OF SQUARES                             STANDARD ERROR OF ESTIMATE
                                        TOTAL          REGR          ERROR                     SE2            SF
                                     1.20000E+01   1.17302E+01   2.69827E-01                6.74568E-02   2.59724E-01
 CONSTANT       3.2718112045E+00

                 3            4            1
B          -8.1329E-01  2.2174E-01  1.0881E-01
SD          1.8746E-01  2.5869E-02  8.0238E-02
B/SD       -4.3384E+00  8.5718E+00  1.3561E+00
PCYI       -9.0815E-01  9.7384E-01  5.6119E-01

INVERSE OF CORRELATION MATRIX (WITH DEPENDENT VARIABLE). SIGN IS REVERSED

    1         -4.447291022E+01

    2         -2.859450413E+01  -2.292247291E+01

    3          6.207627375E+01   4.334414609E+01  -9.136445433E+01

    4          5.905937652E+00   3.701188422E+00  -7.064329810E+00  -2.490325077E+00

INVERSE CHECK

    1          1.000000000E+00

    2         -1.847411113E-13   1.000000000E+00

    3         -1.094235813E-12  -4.689582056E-13   1.000000000F+00

    4         -1.705302566E-13  -1.563194019E-13   3.410605132E-13   1.000000000E+00

COVARIANCE MATRIX OF NORMAL REGRESSION COEFFICIENTS

    1          2.196404947E-02

    2         -1.928872803E-02   2.651622970E-02

    3          5.403291920E-04  -6.629366365E-03   9.590249196E-03

INVERSE OF CROSS-PRODUCT MATRIX (ABOUT MEANS) WITHOUT DEPENDENT VARIABLE

    1          5.209627735E-01

    2         -5.745836161E-02   9.920116953E-03

    3          8.301571555E-03  -1.279172976E-02   9.544196732E-02

             OBSERVED Y           PREDICTED Y       OBS. Y-PRED. Y    (OBS.-PRED.)/OBS.

       1   6.000000000E+00   6.183992064E+00  -1.839920639E-01  -3.066534398E-02
       2   2.000000000E+00   2.271184671E+00  -2.711846708E-01  -1.355923354F-01
       3   3.000000000E+00   2.897144051E+00   1.028559495E-01   3.428531649E-02
       4   5.000000000E+00   5.106301885E+00  -1.063018848E-01  -2.126037696F-02
       5   4.000000000E+00   3.966584869E+00   3.341513079E-02   8.353782697E-03
       6   3.000000000E+00   3.084477628E+00  -8.447762753E-02  -2.815920918E-02
       7   4.000000000E+00   3.710437007E+00   2.895629927E-01   7.239074819F-02
       8   5.000000000E+00   4.779877826E+00   2.201221741E-01   4.402443481F-02

 DURBIN-WATSON STATISTIC =     1.369038858E+00
```

COST ESTIMATE

For the job listed on Sample Input, the central processor time used was 0.269 seconds. At the current rate for the University of Minnesota ($0.20/sec.); the computer time cost $0.05 plus a small charge for output supplies.

Charge to user = computer cost + postage + network overhead
= $0.25 + postage + network overhead

CONTENTS—UMST500

pages	
1– 3	Identification & Abstract
5–13	User Instructions
15–18	I/O
19	Cost—Contents

DESCRIPTIVE TITLE	Invert Ill-Behaved Matrices Using Hotelling and Bodewig's Iterative Technique
CALLING NAME	HRDMIN
INSTALLATION NAME	The University of Iowa University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Robert G. McCord Graduate Student in Comp. Sci. University of Iowa
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/65
PROGRAM AVAILABILITY	Decks and listings presently available
CONTACT	Mrs. Louise R. Levine, Program Librarian, Applications Programming, University Computer Center, The Univ. of Iowa, Iowa City, Ia. 52240 Tel.: (319) 353-5580

FUNCTIONAL ABSTRACT

This routine attempts to invert matrices, particularly ill-behaved matrices such as a Hilbert matrix, to a designated degree of accuracy. Bodewig and Hotelling's method of matrix inversion with iteration is used. The matrix to be inverted and an approximation of its inverse are passed to the routine, and it returns the refined inverse and the product of the original matrix and its inverse. An MXM work area must be passed, where M is the dimension in the main program of the matrix to be inverted. Some matrices may be so ill-conditioned that either an approximation of the inverse cannot be found or the inverse cannot be refined to the degree of accuracy specified. In either case the method cannot be used.

Method

A first approximation to the actual inverse is refined to a desired accuracy using Bodewig and Hotelling's technique.

Let A^{-1} be the true inverse to the matrix A, and let $A_0{}^{-1}$ be a first approximation to A^{-1} such that

continued

$$||I - AA_0^{-1}|| < K < 1$$

where K is a constant, I is the identity matrix, and $||x|| = \max_{1 \leq i \leq n} |x_i|$ is the norm of a matrix X,

$$A_0^{-1} = A^{-1} + E_0$$

where E_0 is an error matrix.

Then $AA_0^{-1} = AA^{-1} + AE_0 = I + AE_0$. Set $R_0 = I - AA_0^{-1}$

$$\begin{array}{ll} A_1^{-1} = A_0^{-1}(I + R_0) & R_1 = I - AA_1^{-1} \\ A_2^{-1} = A_1^{-1}(I + R_1) & R_2 = I - AA_2^{-1} \\ \vdots & \vdots \\ A_p^{-1} = A_{p-1}^{-1}(I + R_{p-1}) & R_p = I - AA_p^{-1} \end{array}$$

The norm of the matrix $A_n^{-1} - A^{-1}$ can be made arbitrarily small by increasing the number of iterations since

$$||A^{-1} - A_p^{-1}|| \leq \frac{n^2P||A_0|| \;\; ||P_0||^{2p}}{1-n \; ||R_0||}$$

where n is the order of A,
p is the number of iterations.

By forcing $||R_0||$ to be less than one, the solution can be made to converge to A^{-1}.

Accuracy

A 24X24 Hilbert matrix was inverted to accuracy of 10^{-4}.

REFERENCES

Householder, A.S., *Principles of Numerical Analysis* (New York: McGraw-Hill Book Co., Inc., 1953), pp. 56-59.

USER INSTRUCTIONS

Usage

The matrices A, B, C, and D must be defined as MXM double precision arrays in the calling program where M is the dimension in the main program of the matrix to be inverted. An approximation to the inverse should be read in or should be calculated by calling a double precision matrix inversion routine (such as the IBM Scientific Subroutine DMINV). Then HRDMIN can be called. On return, IRF should be check to see if refinement to the requested degree of accuracy was possible. If it was not possible, B will contain the approximation passed to the routine. Note that even if there is underflow in the routine which calculates the initial approximation to the inverse, it still may be possible to refine the inverse to the desired accuracy. The returned product A·B should be close to the identity matrix and provides a check on the accuracy.

Calling Sequence

```
CALL HRDMIN(M,N,A,B,C,Y,D,IRF)
```

```
where    M is the dimension in the main program of the matrix
           to be inverted.
         N is the size of the matrix actually used.
         A is the matrix to be inverted (dimension MXM, double
           precision).
         B is an approximation of the inverse.  The refined
           inverse will be returned here (dimension MXM, double
           precision).
         C will contain the product A·B on return (dimension MXM,
           double precision).
         Y is the requested degree of accuracy in double
           precision.
         D is an MXM, double precision work area.
       IRF will contain 0 if refinement is possible
                        1 if refinement to the degree requested
                          is not possible
```

SAMPLE INPUT

```
      0.1D-3      24
0.50000000000000000D 00   0.33333331346511840D 00   0.25000000000000000D 00
0.19999998807907100D 00   0.16666666269302368D 00   0.14285713434219360D 00
0.12500000000000000D 00   0.11111110448837280D 00   0.99999996423721310D-01
0.90909063816070600D-01   0.83333331346511840D-01   0.76923072338104200D-01
0.71428537368774400D-01   0.66666662693023700D-01   0.62500000000000000D-01
0.58823525905609110D-01   0.55555552244186380D-01   0.52631575614213920D-01
0.49999997019767740D-01   0.47619044780731180D-01   0.45454543083906150D-01
0.43478257954120610D-01   0.41666664183139780D-01   0.39999999105930300D-01
0.33333331346511840D 00   0.25000000000000000D 00   0.19999998807907100D 00
0.16666666269302368D 00   0.14285713434219360D 00   0.12500000000000000D 00
0.11111110448837280D 00   0.99999996423721310D-01   0.90909063816070600D-01
0.83333331346511840D-01   0.76923072338104200D-01   0.71428537368774400D-01
0.66666662693023700D-01   0.62500000000000000D-01   0.58823525905609110D-01
0.55555552244186380D-01   0.52631575614213920D-01   0.49999997019767740D-01
0.47619044780731180D-01   0.45454543083906150D-01   0.43478257954120610D-01
0.41666664183139780D-01   0.39999999105930300D-01   0.38461536169052100D-01
0.25000000000000000D 00   0.19999998807907100D 00   0.16666666269302368D 00
0.14285713434219360D 00   0.12500000000000000D 00   0.11111110448837280D 00
0.99999996423721310D-01   0.90909063816070600D-01   0.83333331346511840D-01
0.76923072338104200D-01   0.71428537368774400D-01   0.66666662693023700D-01
0.62500000000000000D-01   0.58823525905609110D-01   0.55555552244186380D-01
0.52631575614213920D-01   0.49999997019767740D-01   0.47619044780731180D-01
0.45454543083906150D-01   0.43478257954120610D-01   0.41666664183139780D-01
0.39999999105930300D-01   0.38461536169052100D-01   0.37037033587694140D-01
0.19999998807907100D 00   0.16666666269302368D 00   0.14285713434219360D 00
0.12500000000000000D 00   0.11111110448837280D 00   0.99999996423721310D-01
0.90909063816070600D-01   0.83333331346511840D-01   0.76923072338104200D-01
0.71428537368774400D-01   0.66666662693023700D-01   0.62500000000000000D-01
0.58823525905609110D-01   0.55555552244186380D-01   0.52631575614213920D-01
0.49999997019767740D-01   0.47619044780731180D-01   0.45454543083906150D-01
0.43478257954120610D-01   0.41666664183139780D-01   0.39999999105930300D-01
0.38461536169052100D-01   0.37037033587694140D-01   0.35714283585548380D-01
0.16666666269302368D 00   0.14285713434219360D 00   0.12500000000000000D 00
0.11111110448837280D 00   0.99999996423721310D-01   0.90909063816070600D-01
0.83333331346511840D-01   0.76923072338104200D-01   0.71428537368774400D-01
0.66666662693023700D-01   0.62500000000000000D-01   0.58823525905609110D-01
0.55555552244186380D-01   0.52631575614213920D-01   0.49999997019767740D-01

          .                         .                         .
          .                         .                         .
          .                         .                         .

0.34482758492231350D-01   0.33333331346511820D-01   0.32258063554763770D-01
0.31249999999999980D-01   0.30303027480840660D-01   0.29411762952804540D-01
0.28571426868438700D-01   0.27777776122093180D-01   0.27027025818824740D-01
0.26315785944461800D-01   0.25641024112701390D-01   0.24999998509883860D-01
0.24390242993831610D-01   0.23809522390365580D-01   0.23255813866853690D-01
0.22727269679307910D-01   0.22222220897674540D-01   0.21739128977060290D-01
0.41666664183139780D-01   0.39999999105930300D-01   0.38461536169052100D-01
0.37037033587694140D-01   0.35714283585548380D-01   0.34482758492231350D-01
0.33333331346511820D-01   0.32258063554763770D-01   0.31249999999999980D-01
0.30303027480840660D-01   0.29411762952804540D-01   0.28571426868438700D-01
0.27777776122093180D-01   0.27027025818824740D-01   0.26315785944461800D-01
0.25641024112701390D-01   0.24999998509883860D-01   0.24390242993831610D-01
0.23809522390365580D-01   0.23255813866853690D-01   0.22727269679307910D-01
0.22222220897674540D-01   0.21739128977060290D-01   0.21276593208312960D-01
0.39999999105930300D-01   0.38461536169052100D-01   0.37037033587694140D-01
0.35714283585548380D-01   0.34482758492231350D-01   0.33333331346511820D-01
0.32258063554763770D-01   0.31249999999999980D-01   0.30303027480840660D-01
0.29411762952804540D-01   0.28571426868438700D-01   0.27777776122093180D-01
0.27027025818824740D-01   0.26315785944461800D-01   0.25641024112701390D-01
0.24999998509883860D-01   0.24390242993831610D-01   0.23809522390365580D-01
0.23255813866853690D-01   0.22727269679307910D-01   0.22222220897674540D-01
0.21739128977060290D-01   0.21276593208312960D-01   0.20833332091569880D-01
/*
```

SAMPLE OUTPUT

```
ORIGINAL MATRIX
   0.50000000000000000D 00  0.33333331346511840D 00  0.25000000000000000D 00
   0.19999998807907100D 00  0.16666662693023680D 00  0.14285713434219360D 00
   0.12500000000000000D 00  0.11111110448837280D 00  0.99999964237213100D-01
   0.90909063816070600D-01  0.83333313465118400D-01  0.76923072338104200D-01
   0.71428537368774400D-01  0.66666662693023700D-01  0.62500000000000000D-01
   0.58823525905609100D-01  0.55555552244186380D-01  0.52631575614213910D-01
   0.49999997019767740D-01  0.47619044780731180D-01  0.45454543083906140D-01
   0.43478257954120600D-01  0.41666664183139780D-01  0.39999999105930300D-01
   0.33333331346511840D 00  0.25000000000000000D 00  0.19999998807907100D 00
   0.16666662693023680D 00  0.14285713434219360D 00  0.12500000000000000D 00
   0.11111110448837280D 00  0.99999964237213100D-01  0.90909063816070600D-01
   0.83333313465118400D-01  0.76923072338104200D-01  0.71428537368774400D-01
   0.66666662693023700D-01  0.62500000000000000D-01  0.58823525905609100D-01
   0.55555552244186380D-01  0.52631575614213910D-01  0.49999997019767740D-01
   0.47619044780731180D-01  0.45454543083906140D-01  0.43478257954120600D-01
   0.41666664183139780D-01  0.39999999105930300D-01  0.38461536169052090D-01
   0.25000000000000000D 00  0.19999998807907100D 00  0.16666662693023680D 00
   0.14285713434219360D 00  0.12500000000000000D 00  0.11111110448837280D 00
   0.99999964237213100D-01  0.90909063816070600D-01  0.83333313465118400D-01
   0.76923072338104200D-01  0.71428537368774400D-01  0.66666662693023700D-01
   0.62500000000000000D-01  0.58823525905609100D-01  0.55555552244186380D-01
   0.52631575614213910D-01  0.49999997019767740D-01  0.47619044780731180D-01
   0.45454543083906140D-01  0.43478257954120600D-01  0.41666664183139780D-01
   0.39999999105930300D-01  0.38461536169052090D-01  0.37037033587694140D-01
   0.19999998807907100D 00  0.16666662693023680D 00  0.14285713434219360D 00
   0.12500000000000000D 00  0.11111110448837280D 00  0.99999964237213100D-01
   0.90909063816070600D-01  0.83333313465118400D-01  0.76923072338104200D-01
   0.71428537368774400D-01  0.66666662693023700D-01  0.62500000000000000D-01
   0.58823525905609100D-01  0.55555552244186380D-01  0.52631575614213910D-01
   0.49999997019767740D-01  0.47619044780731180D-01  0.45454543083906140D-01
   0.43478257954120600D-01  0.41666664183139780D-01  0.39999999105930300D-01
   0.38461536169052090D-01  0.37037033587694140D-01  0.35714283585548370D-01
   0.16666662693023680D 00  0.14285713434219360D 00  0.12500000000000000D 00
   0.11111110448837280D 00  0.99999964237213100D-01  0.90909063816070600D-01
   0.83333313465118400D-01  0.76923072338104200D-01  0.71428537368774400D-01
   0.66666662693023700D-01  0.62500000000000000D-01  0.58823525905609100D-01
   0.55555552244186380D-01  0.52631575614213910D-01  0.49999997019767740D-01
   0.47619044780731180D-01  0.45454543083906140D-01  0.43478257954120600D-01
   0.41666664183139780D-01  0.39999999105930300D-01  0.38461536169052090D-01
   0.37037033587694140D-01  0.35714283585548370D-01  0.34482758492231340D-01
   0.14285713434219360D 00  0.12500000000000000D 00  0.11111110448837280D 00
   0.99999964237213100D-01  0.90909063816070600D-01  0.83333313465118400D-01
   0.76923072338104200D-01  0.71428537368774400D-01  0.66666662693023700D-01
   0.62500000000000000D-01  0.58823525905609100D-01  0.55555552244186380D-01
   0.52631575614213910D-01  0.49999997019767740D-01  0.47619044780731180D-01
   0.45454543083906140D-01  0.43478257954120600D-01  0.41666664183139780D-01
   0.39999999105930300D-01  0.38461536169052090D-01  0.37037033587694140D-01
   0.35714283585548370D-01  0.34482758492231340D-01  0.33333331346511810D-01
   0.12500000000000000D 00  0.11111110448837280D 00  0.99999964237213100D-01
              .                        .                        .
              .                        .                        .
              .                        .                        .

   0.23809522390365580D-01  0.23255813866853690D-01  0.22727269679307910D-01
   0.22222220897674540D-01  0.21739128977060290D-01  0.21276593208312960D-01
   0.39999999105930300D-01  0.38461536169052090D-01  0.37037033587694140D-01
   0.35714283585548370D-01  0.34482758492231340D-01  0.33333331346511810D-01
   0.32258063554763760D-01  0.31249999999999970D-01  0.30303027480840660D-01
   0.29411762552804530D-01  0.28571426868438690D-01  0.27777776122093170D-01
   0.27027025818824740D-01  0.26315785944461800D-01  0.25641024112701390D-01
   0.24999998509883860D-01  0.24390242993831610D-01  0.23809522390365580D-01
   0.23255813866853690D-01  0.22727269679307910D-01  0.22222220897674540D-01
   0.21739128977060290D-01  0.21276593208312960D-01  0.20833332091569880D-01


DETERMINANT=  0.0
```

continued

```
INVERSE MATRIX
   0.12801149401488250D 03  0.55801657195289520D 03 -0.93013874694687650D 04  0.22563933395504730D 05  0.10213614558692180D 04
  -0.47465329619067230D 05  0.10066570339554510D 05  0.48828585373095870D 05  0.25934147442830110D 05  0.64781638329674940D 05
  -0.39827580523157560D 06  0.59062793125375960D 05  0.22801571042622090D 06  0.18214466409274580D 06  0.48358909497821380D 05
  -0.10582115591848310D 06 -0.20190102884468020D 06 -0.15845884472712200D 05 -0.39546161861327570D 05 -0.22628638466053440D 06
   0.23781362189978710D 05  0.50271254823375810D 06  0.12761384345903840D 06 -0.32123240476967580D 06  0.55801653899636410D 03
  -0.24158784849606990D 05  0.16167608367067440D 06 -0.31329345023551060D 06 -0.75617372622554030D 05  0.68646158764880170D 06
  -0.78740396165682380D 05 -0.66095334891432050D 06 -0.28561724575510760D 06 -0.16941018445733030D 07  0.56665923246800610D 07
  -0.46453875957169570D 06 -0.30006304311091060D 07 -0.24672143197329060D 07 -0.85000861450621080D 06  0.12317006458340070D 07
   0.27651427647322710D 07  0.18043040374359120D 06  0.24747513892685070D 06  0.32342338686204330D 07  0.36173444298703580D 06
  -0.69001854947734070D 07 -0.22298694338253470D 07  0.45095451467817400D 07 -0.93013875089083480D 04  0.16167608585522870D 06
  -0.78645655230753130D 06  0.11736668963526830D 07  0.68228592625541860D 06 -0.26392263678263420D 07 -0.48684663635547540D 05
   0.23871364048759070D 07 -0.29291448041487810D 06  0.98579510528114660D 07 -0.22245766359134860D 08  0.83545094210354020D 06
   0.10122895440703390D 08  0.93245785118471360D 07  0.36003832739883300D 07 -0.32675457189565210D 07 -0.11130163748794790D 08
  -0.63274296321392140D 06  0.70956798760592110D 06 -0.11907015664515100D 08 -0.47545890349386820D 07  0.25481482627043040D 03
   0.11000825485344920D 08 -0.17625577878085700D 08  0.22568934130652300D 05 -0.31329345484329830D 06  0.11736669045349370D 07
  -0.10045642952046180D 07 -0.17035730786243150D 07  0.23835544725902200D 07  0.71725996634716050D 06 -0.34405497963370500D 07
   0.85349671869087340D 07 -0.20220573030023470D 08  0.26186354387135650D 08 -0.37378857384414990D 07 -0.71926186550122980D 07
  -0.11549066489832660D 08 -0.17449804805849460D 07 -0.30305252317689250D 06  0.16700260047146900D 08  0.19453051362418590D 07
  -0.33568903466793300D 07  0.78723334260605090D 07  0.98823721789719250D 07 -0.24641480293844100D 08 -0.17617964961329990D 08
   0.21409776058331840D 08  0.10213613931852820D 04 -0.75617372527578550D 05  0.63228592651986570D 06 -0.17035730660440050D 07
  -0.71948133414006790D 05  0.34237225653804030D 07 -0.12143728036663480D 07  0.73171127406615380D 07 -0.21722825168281680D 08
   0.58968672398786040D 07  0.69797370593100910D 07  0.15727937151025080D 08 -0.12823372825369770D 08  0.19809330074033190D 07
  -0.12767527543718260D 08  0.96204647346944820D 07 -0.34958838050251870D 07 -0.34815638036657350D 07 -0.57018525947928280D 07
   0.20282439714912590D 08  0.88954198292480800D 07 -0.22304569377171510D 08 -0.17520647950062020D 07  0.63082126376355380D 07
  -0.47465329552059430D 05  0.68646159136449130D 06 -0.26392263676581410D 07  0.23835544625757300D 07  0.34237225658956760D 07
  -0.62561245999719480D 07  0.85131028141982110D 07 -0.16465383753065950D 08  0.86559607427579750D 07  0.15285784460782630D 08
  -0.27119487843267960D 08 -0.53600599738585290D 07  0.26731399480223890D 08  0.72134497225728250D 07  0.14317837416215930D 08
  -0.15329635089552830D 08 -0.25694660814632410D 08 -0.74315287380822470D 07  0.15365510452210450D 08 -0.15253907568384070D 08
  -0.18199435825682420D 08  0.52731024696184710D 08  0.19144419880618890D 08 -0.34664642936526700D 08  0.10064570638359660D 05

              .                                     .                                     .
              .                                     .                                     .
              .                                     .                                     .

   0.16670375406653650D 09 -0.19064700100287260D 09 -0.78814501458528770D 08 -0.99534715296215520D 08  0.98782090981383670D 08
   0.80122883704919370D 08 -0.44592189332775220D 08  0.10180263944954290D 08  0.12444916370835730D 09  0.69230727766568360D 08
  -0.15355672929742560D 09 -0.73531255549079680D 08 -0.30835190681999030D 08  0.84779195515672010D 08  0.50271255308152340D 06
  -0.69001856136178910D 07  0.25461482834048930D 08 -0.24641480522714310D 08 -0.22304569707659390D 08  0.52731025290468770D 08
  -0.31284257502863240D 08 -0.52933244032852100D 08  0.19323963434120550D 08  0.19340790535905430D 08  0.41247478012970590D 09
  -0.14242853681456970D 09 -0.43976341401044130D 09 -0.16795448396282690D 09 -0.48713891160354720D 07  0.10519540731950800D 09
   0.30446690620183660D 09  0.14786440054461310D 09  0.11369197960595490D 09  0.19365823909493590D 09 -0.73531256626986070D 08
  -0.78279248426769060D 09 -0.12925806799119010D 09  0.48399082656112780D 09  0.12761384384716340D 06 -0.22298694505783550D 07
   0.11000825475903740D 08 -0.17617964965262180D 08 -0.17520647874610710D 07  0.19144419859262170D 08 -0.12557851784223560D 07
   0.23364328600122700D 07 -0.87253820062610730D 07 -0.21858619435854370D 08  0.13596658136607650D 08  0.83046173508710540D 08
  -0.14898993202981600D 09  0.11594528359059790D 09 -0.16240291783778690D 09  0.12479320599052770D 09 -0.15261790037470240D 08
   0.10762442234694580D 09 -0.11691251785483620D 09  0.11113540454240100D 09 -0.30835191156617380D 08 -0.12925806719296410D 09
  -0.54284986980421210D 07  0.73806250931234010D 08 -0.32123240534347300D 06  0.45095451814700030D 07 -0.17625577817080630D 08
   0.21409775983984280D 08  0.63082127077297210D 07 -0.34666430265971400D 08  0.25255162728397280D 09  0.34495735631126950D 08
  -0.36861831278563500D 07 -0.14119344262394900D 08 -0.26509094852461460D 09  0.41226550826079690D 08  0.34274573718368100D 09
   0.45594670981140960D 08  0.10394338647890560D 09 -0.16389498000309920D 09 -0.13034740666437000D 09 -0.13387964165463010D 09
  -0.62035872918277130D 07 -0.20426015725969880D 09  0.84779195851068130D 08  0.48339082360824800D 09  0.73806250637406240D 08
  -0.29400230141193940D 09
```

continued

```
IDENTITY MATRIX
  0.1000000000109140D 01 -0.17462298274040220D-09  0.80617610365152360D-08  0.10913936421275140D-08  0.58739737522602080D-08
  0.25611370801925660D-08  0.40745362639427180D-09 -0.69849193096160890D-08  0.88475644588470460D-08  0.10535585625337600D-07
  0.74272975325584410D-07 -0.16763806343078610D-07  0.22584572434425350D-07 -0.12107193470001220D-07 -0.30267983675003050D-08
  0.13737007975578310D-07 -0.16298145055770870D-07  0.27939677238464360D-08  0.11277734301984310D-07  0.0
  0.11641532182693480D-08 -0.55879354476928710D-07 -0.27939677238464360D-08  0.27474015951156620D-07  0.45474735088646410D-11
  0.9999999981082510D 00  0.77416189014911650D-08  0.17025740817189220D-08  0.58644218370318410D-08  0.32596290111541750D-08
 -0.71304384618997580D-09 -0.62864273786544800D-08  0.73632691055536270D-08  0.78871380527749340D-08  0.67268056015968320D-07
 -0.15832483768463130D-07  0.35390257835388180D-07 -0.12107193470001220D-07 -0.13969838619232180D-08  0.13737007975578310D-07
 -0.12805685400962830D-07  0.23283064365386960D-08  0.10040821507573130D-07 -0.69849193096160890D-09  0.23283064365386960D-08
 -0.74505805969238280D-07 -0.27939677238464360D-08  0.12340024113655090D-07  0.63664629124104970D-11 -0.11641532182693480D-09
  0.1000000068248480D 01  0.16298145055770870D-08  0.54424162954092030D-08  0.23283064365386960D-08 -0.10040821507573130D-08
 -0.62864273786544800D-08  0.59808371588587760D-08  0.56752469390630720D-08  0.45401975512504580D-07 -0.15832483768463130D-07
  0.30966475605964660D-07 -0.10477378964424130D-07 -0.16298145055770870D-08  0.12805685400962830D-07 -0.13969838619232180D-07
  0.34924596548080440D-08  0.82072801887989040D-08 -0.23283064365386960D-09  0.18626451492309570D-08 -0.37252902984619140D-07
 -0.46566128730773930D-09  0.69849193096160890D-08  0.0                     -0.10186340659856800D-09  0.65920175984501840D-08
  0.1000000013533280D 01  0.50931703299283980D-08  0.25611370801925660D-08 -0.23283064365386960D-09 -0.51222741603851320D-08
  0.45693013817071910D-08  0.73487171903252600D-08  0.44470652937889100D-07 -0.15366822481155400D-07  0.19557774066925050D-07
 -0.95460563898086550D-08 -0.11641532182693480D-08  0.13737007975578310D-07 -0.14202669262886050D-07  0.30267983675003050D-08
  0.89203240349888800D-08 -0.23283064365386960D-09  0.27939677238464360D-08 -0.37252902984619140D-07 -0.11641532182693480D-08
  0.93132257461547850D-08  0.27284841053187840D-11 -0.13096723705530170D-09  0.62136678025126460D-08  0.13969838619232180D-08
  0.1000000004685717D 01  0.16298145055770870D-08 -0.18044374883174900D-08 -0.41909515857696530D-08  0.49185473471879960D-08
  0.59662852436304090D-08  0.41909515857696530D-07 -0.15832483768463130D-07  0.25145709514617920D-07 -0.90803951025009150D-08
  0.0                      0.12340024113655090D-07 -0.13504177331924440D-07  0.20954757928848270D-08  0.69994712248444560D-08
  0.0                      0.13969838619232180D-08 -0.37951394915580750D-07 -0.11641532182693480D-08  0.19092112779617310D-07
  0.45474735088646410D-11 -0.16007106751203540D-09  0.59808371588587760D-08  0.17025740817189220D-08  0.43510226532816890D-08
  0.1000000013969840D 01 -0.96042640507221230D-09 -0.41909515857696530D-08  0.53260009735822680D-08  0.59953890740871430D-08
  0.40279701352119450D-07 -0.14901161193847660D-07  0.23748725652694700D-07 -0.83819031715393070D-08 -0.20954757928848270D-08
  0.10710209608078000D-07 -0.12805685400962830D-07  0.23283064365386960D-08  0.70285750553011890D-08  0.0
 -0.23283064365386960D-09 -0.27474015951156620D-07 -0.69849193096160890D-09  0.12340024113655090D-07  0.45474735088646410D-11
 -0.20372681319713590D-09  0.56170392781496050D-08  0.15570549294352530D-08  0.43073669075965880D-08  0.16298145055770870D-08

            .                                  .                                                            .
            .                                  .                                                            .
            .                                  .                                                            .

 -0.46566128730773930D-09  0.81490725278854370D-08 -0.74505805969238280D-08  0.46566128730773930D-09  0.37689460441470150D-08
  0.1000000006984920D 01 -0.23283064365386960D-09 -0.22817403078079220D-07  0.46566128730773930D-09  0.44237822294235230D-08
  0.36379788070917130D-11 -0.10186340659856800D-09  0.33178366720676420D-08  0.10913936421275140D-08  0.25174813345074650D-08
  0.14697434380650520D-08 -0.15861587598919870D-08 -0.32741809263825420D-08  0.19203528101444240D-08  0.25465851649641990D-08
  0.21653249859809880D-07 -0.94150891527533530D-08  0.13737007975578310D-07 -0.57916622608900070D-08 -0.93132257461547860D-09
  0.79162418842315670D-08 -0.69849193096160890D-08  0.16298145055770870D-08  0.37980498746037480D-08  0.0
  0.9999999860301610D 00 -0.21886080503463740D-07 -0.46566128730773930D-09  0.46566128730773930D-08  0.45474735088646410D-11
 -0.87311491370201110D-10  0.31432136893272400D-08  0.10913936421275140D-08  0.24447217583656310D-08  0.11496013030409810D-08
 -0.15425030142068860D-08 -0.28958311304450030D-08  0.18771970644593240D-08  0.34051481634378430D-08  0.21187588572502140D-07
 -0.88039087131619450D-08  0.13271346688270570D-07 -0.40745362639427180D-08  0.0                      0.72177499532699550D-08
 -0.67520886659622190D-08  0.23283064365386960D-08  0.42637111619114880D-08  0.69849193096160890D-09 -0.69849193096160890D-09
  0.9999999804422259D 00 -0.23283064365386960D-09  0.34924596548080440D-08  0.27284841053187840D-11 -0.10186340659856800D-09
  0.30559021979570390D-08  0.10622898116707800D-08  0.23428583517670630D-08  0.11737051334977150D-08 -0.15133991837501530D-08
 -0.29249349609017370D-08  0.17462298274040220D-08  0.26338966563344000D-08  0.20256265997888660D-07 -0.86001813499648050D-08
  0.13504177331924440D-07 -0.47875801101326940D-08  0.0                      0.72177499532699580D-08 -0.60535967350006100D-08
  0.69849193096160890D-09  0.35506673157215120D-08  0.11641532182693480D-08 -0.23283064365386960D-09 -0.25611370801925660D-07
  0.1000000006984920D 01  0.39581209421157840D-08  0.27284841053187840D-11 -0.72759576141834260D-10  0.29976945370435710D-08
  0.10477378964424130D-08  0.21827872842550280D-08  0.12660166248679160D-08 -0.14697434380650520D-08 -0.25902409106453900D-08
  0.17753336578607560D-08  0.29831426218152050D-08  0.20256265997888660D-07 -0.80181052908301350D-08  0.12107193470001220D-07
 -0.42637111619114880D-08 -0.16298145055770870D-08  0.69849193096160890D-08 -0.65192580223083500D-08  0.20954757928848270D-08
  0.47293724492192270D-08  0.23283064365386960D-09  0.0                     -0.20721927285194400D-07 -0.23283064365386960D-09
  0.1000000041909520D 01
```

COST ESTIMATE

The costing algorithm is based on the amount of computing time, core storage, I/O accesses, cards in and out, and lines and pages printed. The total cost for the Sample Output above was $1.86.

Charge to user = computer cost + postage + network overhead
= $1.86 + postage + network overhead

CONTENTS—HRDMIN

pages

1– 2 Identification & Abstract
3 User Instructions
5– 8 I/O
9 Cost—Contents

DESCRIPTIVE TITLE	Biomedical Computer Programs
CALLING NAME	BMD
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	Health Sciences Computing Facility University of California, Los Angeles
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from authors.
CONTACT	Dr. C. B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

The BMD package contains a variety of statistical programs in the following areas.

Description and Tabulation

Multivariate Analysis
a. Factor Analysis
b. Discriminant Analysis
c. Canonical Analysis

Regression Analysis
a. Linear
b. Polynomial
c. Asymptotic

Analysis of Variance and Covariance

Time Series Analysis

Special Programs
a. Life Table and Survival Rate
b. Contingency Table Analysis
c. Biological Assay
d. Guttman Scaling

continued

REFERENCES

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs* (Berkeley: Univ. of Cal. Press, 1968).

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs, X-series Supplement* (Berkeley: Univ. of Cal. Press, 1969).

USER INSTRUCTIONS

The BMD manuals referenced below contain a description of each program, including sample input and output.

Most of the sixty programs in this package require a minimum partition size of 150K.

Deck Setup

```
Standard JOB Card
//     EXEC  BMD,PROGRAM=BMDXXX
       where XXX is the program code name, such as 01D
//GO.SYSIN    DD    *

       (BMD Program Control Cards)
       (Data)

/*
```

REFERENCES

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs* (Berkeley: Univ. of Cal. Press, 1968).

Dixon, W.J., Ed., *BMD: Biomedical Computer Programs, X-series Supplement* (Berkeley: Univ. of Cal. Press, 1969).

SAMPLE INPUT and SAMPLE OUTPUT

Interested persons should consult the references listed under the User Instructions.

COST ESTIMATE

Interested persons should consult the contact person.

Charge to user = S/360 charges + handling/consultation charges + postage + network overhead

CONTENTS—BMD

pages

1– 2 Identification & Abstract
3 User Instructions
5 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	General Purpose Simulation System
CALLING NAME	GPSS
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	360 Assembly Language
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use at Washington University. Available for distribution from IBM.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

GPSS is a transaction-oriented language designed for conducting evaluations and experiments concerning the behavior of systems, methods and processes. It has a modular structure which permits "transactions" to flow through the system, where their interactions can be observed and modified. A "clock" is maintained by which events are either scheduled to occur or else determined by one of the eight random number generators provided. Information can be obtained regarding sequencing of operations, scheduling and allocation rules, inventories, queuing disciplines, machine failures, etc. In general, various trade-offs between cost and performance can be studied.

REFERENCES

General Purpose Simulation System/360, Introductory User's Manual (IBM Manual GH20-0304-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, User's Manual (IBM Manual GH20-0326-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

General Purpose Simulation System/360, OS (360A-CS-17X) Operator's Manual (IBM Manual H20-0311), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

USER INSTRUCTIONS

General Deck Setup

```
JOB Card
//  EXEC    GPSSGO
//GPSS.SYSIN    DD   *
   (GPSS Model)
/*
```

The above procedure requires a minimum partition of 90K and does not permit the use of the specialized features: UPDATE, WRITE-JOBTAPE, or READ-SAVE. The setup for these features is described in Ref. 3, and will increase the partition requirements beyond 100K.

REFERENCES

1. *General Purpose Simulation System/360, Introductory User's Manual* (IBM Manual GH20-0304-0), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

2. *General Purpose Simulation System/360, User's Manual* (IBM Manual GH20-0326-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

3. *General Purpose Simulation System/360, OS (360A-CS-17X) Operator's Manual* (IBM Manual H20-0311), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1969).

516

SAMPLE INPUT

```
  JOB Card
//  EXEC    GPSSGO
//GPSS.SYSIN    DD   *
BLOCK
NUMBER  *LOC   OPERATION  A,B,C,D,E,F,G              COMMENTS
        *      SIMULATE
         5     FUNCTION   RN2,C24            EXPONENTIAL
        0      0     .1   .104  .2    .222  .3    .355  .4    .509  .5    .69
        .6     .915  .7   1.2   .75   1.38  .8    1.6   .84   1.83  .88   2.12
        .9     2.3   .92  2.52  .94   2.81  .95   2.99  .96   3.2   .97   3.5
        .98    3.9   .99  4.6   .995  5.3   .998  6.2   .999  7     .9997 8
         1     FUNCTION   RN3,D3             TYPE
        .42    1     .62  2 1.0 3
1              GENERATE   5,FN5
2              ASSIGN     1,FN1
3              QUEUE      P1
4              ENTER      P1
         1     STORAGE    2                  REGULAR GAS
         2     STORAGE    1                  HI TEST GAS
         3     STORAGE    2                  FUEL OIL
5              DEPART     P1
6              ADVANCE    19,5
7              LEAVE      P1
8              TERMINATE  1
               START      2000
               REPORT
               EJECT
        STO    TITLE      ,S1=REGULAR GAS.....S2=HIGH TEST......S3=FUEL OIL
               SPACE      3
               SPACE      3
        QUE    TITLE      ,Q1=REGULAR GAS......Q2=HIGH TEST......Q3=FUEL OIL
               EJECT
               END
         5     FUNCTION   RN2   C24
        0      0     .1   .104  .2    .222  .3    .355  .4    .509  .5    .69
        .6     .915  .7   1.2   .75   1.38  .8    1.6   .84   1.83  .88   2.12
        .9     2.3   .92  2.52  .94   2.81  .95   2.99  .96   3.2   .97   3.5
        .98    3.9   .99  4.6   .995  5.3   .998  6.2   .999  7     .9997 8
         1     FUNCTION   RN3   D3
        .42    1     .62  2     1.0   3
         1     GENERATE   5     FN5
         2     ASSIGN     1     FN1
         3     QUEUE      P1
         4     ENTER      P1
         1     STORAGE    2
         2     STORAGE    1
         3     STORAGE    2
         5     DEPART     P1
         6     ADVANCE    19
         7     LEAVE      P1
         8     TERMINATE  1
               START      2000
/*
```

SAMPLE OUTPUT

```
S1=REGUALR GAS.....S2=HIGH TEST......S3=FUEL OIL
 STORAGE       CAPACITY


S1=REGULAR GAS.....S2=HIGH TEST......S3=FUEL OIL

 STORAGE       CAPACITY      AVERAGE        AVERAGE        ENTRIES          AVERAGE          CURRENT
                             CONTENTS     UTILIZATION                     TIME/TRAN         CONTENTS
       1             2         1.701          .850             807          19.381                1
       2             1          .799          .799             390          18.848                1
       3             2         1.671          .835             807          19.029                2



Q1=REGULAR GAS......Q2=HIGH TEST......Q3=FUEL OIL

 QUEUE       MAXIMUM     AVERAGE      TOTAL       ZERO   PERCENT        AVERAGE     $AVERAGE          TABLE
            CONTENTS    CONTENTS     ENTRIES    ENTRIES   ZEROS       TIME/TRANS   TIME/TRANS         NUMBER
     1            24       3.620        816        167     20.4          40.772       51.263
     2            15       2.501        399         82     20.5          57.621       72.526
     3            10       1.562        807        221     27.3          17.796       24.508
 $AVERAGE TIME/TRANS = AVERAGE TIME/TRANS EXCLUDING ZERO ENTRIES
```

COST ESTIMATE

The above sample problem, executed in class A (100K high speed core), required 36 seconds of central processor unit time at a cost of $5.04.

Charge to user = S/360 charges + handling/consultation charges + postage + network overhead

= $5.04 + handling/consultation charges + postage + network overhead

CONTENTS—GPSS

pages	
1	Identification & Abstract
3	User Instructions
5– 6	I/O
7	Cost—Contents

000 0134

DESCRIPTIVE TITLE	Continuous System Modeling Program
CALLING NAME	CSMP
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	IBM Application Program
LANGUAGE	CSMP imbedded in FORTRAN
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Proprietary; available for use but not for distribution.
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

CSMP performs a simulation of a continuous system, obtaining solutions to problems expressed in the form of systems of differential equations or analog block diagrams. Typical applications might be a control engineer's study of the effectiveness of various control system designs, or a study of a cardiovascular system model.

CSMP provides a basic set of functional blocks with which the components of a continuous system may be represented, and it accepts application-oriented statements for defining the connections between these functional blocks. It also accepts FORTRAN IV statements which can be used to handle non-linear and time-variant problems. Input and output are facilitated by means of user-oriented control statements.

REFERENCES

System/360 Continuous System Modeling Program (360A-CX-16X) User's Manual (IBM Manual GH20-0367-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

continued

000 0134

System/360 Continuous System Modeling Program (360A-CX-16X) Operator's Manual (IBM Manual H20-0368), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

System/360 Continuous System Modeling Program (360A-CX-16X) Application Description (IBM Manual GH20-0240-2), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

USER INSTRUCTIONS

Fixed format output is provided for printing and plotting at selected increments of the independent variable. Output may also be controlled through user supplied formats. The user's manual describes diagnostic error messages. Minimum core storage required is 90K bytes.

General Deck Setup

```
JOB Card
//  EXEC    CSMPGO
//CSMP.SYSIN  DD  *

    (CSMP Model)

/*
```

Further information can be obtained from the references listed below.

REFERENCES

System/360 Continuous System Modeling Program (360A-CX-16X) User's Manual (IBM Manual GH20-0367-3), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

System/360 Continuous System Modeling Program (360A-CX-16X) Operator's Manual (IBM Manual H20-0368), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

System/360 Continuous System Modeling Program (360A-CX-16X) Application Description (IBM Manual GH20-0240-2), (White Plains, N.Y.: IBM Corp. Tech. Pub. Depart., 1968).

SAMPLE INPUT

```
TITLE              CABLE REEL CONTROL DESIGN
*                        APRIL 1, 1968
INITIAL
                K1 = (D**2)/(2.0*PI*W)
PARAMETER        D = 0.1       ,        W = 2.0
CONSTANT        PI = 3.14159
PARAMETER    RFULL = 4.0       ,  REMPTY = 2.0
*
DYNAMIC

                 I = 18.5 * (R**4)   - 221.0
            TH2DOT = TORQUE / I
            TH1DOT = INTGRL(0.0 , TH2DOT )
                 R = INTGRL(RFULL , (-K1 * TH1DOT))
             ERROR = VDESIR - VM
PARAMETER   VDESIR = 50.0
             CONTL = GAIN * ERROR
*     (REMEMBER TO ADD CONTROL LIMITING ON NEXT TRY)
PARAMETER     GAIN = 0.5
            TORQUE = 500.0 * DUMMY
             DUMMY = REALPL(0.0, 1.0, CONTL)
              VACT = R * TH1DOT
                VM = REALPL(0.0, 0.5, VACT)
FINISH           R = 2.0
TIMER  DELT =.05, FINTIM = 20.0, PRDEL=0.5,OUTDEL=0.5
PRINT    VACT,VM,ERROR,CONTL  ,TORQUE,R,I
PRTPLT   VACT
LABEL       PRELIMINARY TEST OF SYSTEM STABILITY (GAIN = 0.5)
METHOD   RECT
END
PARAMETER     GAIN = 1.5
RESET LABLE
LABEL       PRELIMINARY TEST OF SYSTEM STABILITY (GAIN = 1.5)
END STOP
```

/

SAMPLE OUTPUT

```
                   *** CSMP/360 SIMULATION DATA ***

TITLE              CABLE REEL CONTROL DESIGN

PARAMETER        D = 0.1     ,       W = 2.0

CONSTANT        PI = 3.14159

PARAMETER    RFULL = 4.0     ,  REMPTY = 2.0

PARAMETER   VDESIR = 50.0

PARAMETER     GAIN = 0.5

FINISH           R = 2.0

TIMER  DELT =.05, FINTIM = 20.0, PRDEL=0.5,OUTDEL=0.5

PRINT    VACT,VM,ERROR,CONTL  ,TORQUE,R,I

PRTPLT   VACT

LABEL       PRELIMINARY TEST OF SYSTEM STABILITY (GAIN = 0.5)

METHOD   RECT

END
```

```
TIMER VARIABLES
 DELT  = 5.0000E-02
 DELMIN= 2.0000E-06
 FINTIM= 2.0000E 01
 PRDEL = 5.0000E-01
 OUTDEL= 5.0000E-01
```

```
           CABLE REEL CONTROL DESIGN                              RECT      INTEGRATION

 TIME         VACT        VM          ERROR       CONTL       TORQUE      R           I
0.0          0.0         0.0         5.0000E 01  2.5000E 01  0.0         4.0000E 00  4.5150E 03
5.0000E-01   1.0931E 00  2.5617E-01  4.9744E 01  2.4872E 01  5.0101E 03  4.0000E 00  4.5148E 03
1.0000E 00   3.9555E 00  1.7263E 00  4.8274E 01  2.4137E 01  7.9291E 03  3.9997E 00  4.5138E 03
1.5000E 00   7.8155E 00  4.4856E 00  4.5514E 01  2.2757E 01  9.4633E 03  3.9992E 00  4.5111E 03
2.0000E 00   1.2156E 01  8.1525E 00  4.1847E 01  2.0924E 01  1.0057E 04  3.9982E 00  4.5065E 03
2.5000E 00   1.6624E 01  1.2318E 01  3.7681E 01  1.8841E 01  1.0016E 04  3.9968E 00  4.4998E 03
    .                        .                       .                       .
    .                        .                       .                       .
    .                        .                       .                       .
9.0000E 00   4.7452E 01  4.6669E 01  3.3313E 00  1.6657E 00  1.4524E 03  3.9493E 00  4.2793E 03
9.5000E 00   4.8004E 01  4.7378E 01  2.6216E 00  1.3108E 00  1.1674E 03  3.9445E 00  4.2574E 03
1.0000E 01   4.8435E 01  4.7942E 01  2.0581E 00  1.0290E 00  9.3300E 02  3.9396E 00  4.2353E 03
1.0500E 01   4.8767E 01  4.8384E 01  1.6162E 00  8.0811E-01  7.4233E 02  3.9347E 00  4.2131E 03
    .                        .                       .                       .
    .                        .                       .                       .
    .                        .                       .                       .
1.9000E 01   4.9511E 01  4.9520E 01  4.8030E-01  2.4015E-01  1.1583E 02  3.8486E 00  3.8378E 03
1.9500E 01   4.9504E 01  4.9511E 01  4.8869E-01  2.4435E-01  1.1795E 02  3.8435E 00  3.8168E 03
2.0000E 01   4.9497E 01  4.9504E 01  4.9599E-01  2.4799E-01  1.2002E 02  3.8384E 00  3.7947E 03
```

COST ESTIMATE

The sample problem was taken from Ref. 3, User Instructions. It was executed in class A (100K, high speed core) and required 0.5 minutes of central processor unit time at a cost of $4.20.

Charge to user = S/360 charges + handling/consultation charges + postage + network overhead

= $4.20 + handling/consultation charges + postage + network overhead

CONTENTS—CSMP

pages	
1– 2	Identification & Abstract
3	User Instructions
5– 6	I/O
7	Cost—Contents

DESCRIPTIVE TITLE	A Primal-Dual Transportation Algorithm
CALLING NAME	TRANSPRT
INSTALLATION NAME	Washington University Computing Facilities
AUTHOR(S) AND AFFILIATION(S)	D.E. Burlingame Washington University Computing Facilities
LANGUAGE	FORTRAN IV
COMPUTER	IBM 360/50
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. C.B. Drebes, Mgr., Scientific Data Processing, Computing Facilities, Box 1098, Washington University, St. Louis, Mo. 63130 Tel.: (314) 863-0100 ext. 3141

FUNCTIONAL ABSTRACT

TRANSPRT solves a standard transportation problem using the primal-dual transportation algorithm. This program solves a model in which the objective is to "transport" a single commodity from various origins to different destinations at a minimum total shipping cost. The availability at each origin, the demand at each destination, and the cost to ship one unit of the project from any origin to any destination are required inputs to this model. This model can be applied to certain other types of industrial or business problems that have nothing to do with shipping. Personnel assignment, machine assignment, product and inventory scheduling are a few such applications.

The program utilizes integer arithmetic and requires integer input. If the sum of the origin availabilities is not equal to the sum of the destination requirements, an artificial origin or destination is set up, with zero costs, to handle the excess. The program will handle up to 50 origins and 150 destinations, using 94K bytes of core storage.

REFERENCES

Hadley, G., *Linear Programming* (Reading, Mass: Addison-Wesley Pub. Co., Inc., 1962).

USER INSTRUCTIONS

Deck Setup

```
Standard JOB Card
//    EXEC FORTRAN,LIBRARY='DS09816.LINLIB',PROG[illegible]=TRANSPRT
//GO.SYSIN    DD    *
Program Parameter Card          ┐
Origin Availability Cards       │ may be repeated for
Destination Requirement Cards   │ additional problems
Cost Coefficient Cards          ┘
/*
```

Program Control and Data Cards

Program Parameter Card

This card defines the number of origins and destinations, along with a punched option.

Columns	*Contents*
1– 5	Number of origins $M \leq 50$; integer value, right-justified.
6–10	Number of destinations $N \leq 150$; integer value, right-justified.
16–19	PUNC: punched output of the optimal solution is desired; otherwise leave blank.

Origin Availabilities and Destination Requirements Card Sets

These integer values are read in by an 8I10 FORTRAN format, with each set beginning on a new card.

Cost Coefficient Cards

For each origin the costs are read by an 8I10 FORTRAN format. The costs for each successive origin must begin on a new card.

Example: (M = 2, N = 9)

first card: c_{11} c_{12} c_{13} c_{14} c_{15} c_{16} c_{17} c_{18}

second card: c_{19}

third card: c_{21} c_{22} c_{23} c_{24} c_{25} c_{26} c_{27} c_{28}

fourth card: c_{29}

SAMPLE INPUT and SAMPLE OUTPUT

Test Problem

Origins	*Availabilities*	*Destinations*	*Requirements*
1	50	1	30
2	40	2	50
3	60	3	20
4	31	4	40
		5	30
		6	11

Shipping Costs

	Destination					
Origin	1	2	3	4	5	6
1	2	1	3	3	2	5
2	3	2	2	4	3	4
3	3	5	4	2	4	1
4	4	2	2	1	2	2

Optimal Solution

Minimum cost = 330.

x_{ij} = amount shipped from origin i to destination j

$x_{12} = 30$, $x_{22} = 20$, $x_{31} = 30$, $x_{44} = 21$

$x_{15} = 20$, $x_{23} = 20$, $x_{34} = 19$, $x_{45} = 10$

$x_{36} = 11$

COST ESTIMATE

The above sample problem was executed in class A (100K high speed core) and required 0.02 minutes of central processor unit time at a cost of $0.17.

Charge to user = S/360 charges + handling/consultation charges + postage + network overhead

= $0.17 + handling/consultation charges + postage + network overhead

CONTENTS—TRANSPRT

pages
1 Identification & Abstract
3 User Instructions
5 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Catholic University Computer Center
INSTALLATION NAME	The Catholic University of America Computer Center
LANGUAGES	FORTRAN IV BASIC COBOL ASSEMBLER (MACRO) and various special languages and utilities
COMPUTER	PDP-10 with 64K words of core
AVAILABILITY	Remote access teletype terminals and Batch processing
CONTACT	Dr. Andrew G. Favret, Dir., Computer Center, The Catholic Univ. of America, Washington, D.C. 20017 Tel.: (202) 529-6000 ext. 661

FUNCTIONAL ABSTRACT

The PDP-10 is a timesharing system with a rather flexible Monitor. The equipment includes: two 7-track tape drives, one RPO2 disk pack drive, eight DECtapes, and one CalComp on-line graphic plotter, in addition to card reader, line printer and paper-tape reader/punch. All of the peripherals are available directly or indirectly to remote users. Remote access low-speed terminals such as Teletype models 33 and 35 should be used (ASCII Code). The dial-up number is (202) 526-3300. It is necessary to obtain project/programmer numbers and password before using the system.

USER INSTRUCTIONS

The system is normally available 7 days a week, 24 hours a day, except for scheduled maintenance. However, operators are normally on duty

Monday — Friday	8 a.m. – 10 p.m.
Saturday	9 a.m. – 5 p.m.

Some recommended references are listed below.

REFERENCES

PDP–10 User's Guide (Washington, D.C.: Catholic Univ. Comp. Ctr., March 1971). Available from Catholic University on request.

PDP–10 Timesharing Handbook (Maynard, Mass: Digital Equip. Corp., 1970). Details interaction with the timesharing monitor from programming and operation view-point; includes sections on BASIC, AID, FORTRAN and utility programs such as LINED and PIP. ($5.00)

PDP–10 Reference Handbook (Maynard, Mass.: Digital Equip. Corp.). Describes the PDP–10 central processor and instruction repertoire, word formats, memory characteristics and assembler conventions. ($5.00)

COST ESTIMATE

Current charges for timesharing use are the sum of the following.

Description	*Cost*
Central Processor Time	$150/hr
Kilo—core Seconds	$20/hr before 5 p.m.
Connect Time	$2.40/hr

Charge to user = computer costs + communications + network overhead

CONTENTS

pages
1 Identification & Abstract
3 User Instructions
5 Cost—Contents

DESCRIPTIVE TITLE	Chi Square
CALLING NAME	UMST520
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6.)
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program computes the chi-square criterion on two-way frequency tables. Input data consist of observations in a contingency table, commonly referred to as an r X c table, with the following restrictions:

$$r \geq 2, \quad c \geq 2$$

where r is the number of rows and c is the number of columns. Row and column sums and expected frequencies may be selected as an optional output. Yates' continuity correction is applied in the case of 2 X 2 tables.

Formulae

Let x_{ij} be the data element in the ith row and jth column.

Sums of columns: $n_{.j} = \sum_{i=1}^{r} x_{ij}$ where r is the number of rows.

Sums of rows: $n_{i.} = \sum_{j=1}^{c} x_{ij}$ where c is the number of columns.

continued

Total: $N = \sum_{j=1}^{c} n_{.j}$

Chi square:

$$\text{Chi square} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(x_{ij} - \frac{n_{.j}n_{i.}}{N}\right)^2}{\frac{n_{.j}n_{i.}}{N}}$$

$$= N \left[\sum_{i=1}^{r} \sum_{j=1}^{c} \left(\frac{x_{ij}^{2}}{n_{i.}n_{.j}}\right) - 1 \right]$$

The latter form is used in the program.

Degrees of freedom: d.f. = (r–1) (c–1)

Yates' continuity corrections:

Given the 2 X 2 table:

	a	b	a+b
	c	d	c+d
total	a+c	b+d	r

Use this table if (a+c)(a+b)/r = a

If (a+c)(a+b)/r > a, then use

$$\boxed{\begin{array}{ll} a + 1/2 & b - 1/2 \\ c - 1/2 & d + 1/2 \end{array}} \quad \text{instead of} \quad \boxed{\begin{array}{ll} a & b \\ c & d \end{array}}$$

If (a+c)(a+b)/r < a, then use

$$\boxed{\begin{array}{ll} a - 1/2 & b + 1/2 \\ c + 1/2 & d - 1/2 \end{array}} \quad \text{instead of} \quad \boxed{\begin{array}{ll} a & b \\ c & d \end{array}}$$

Expected frequencies: $f_{ij} = n_{.j}n_{i.}/N$

USER INSTRUCTIONS

System Control Cards (will be prepared by University of Minnesota personnel)

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8–10	NCHI	Number of sets of chi-square data. NCHI > 0
11–13	NR	Number of rows, r
14–16	NC	Number of columns, c
17		blank
18	NFC	Number of Format Cards
19–20	MS	1: marginal sums are to be printed 0: otherwise
21–22	NF	1: expected frequencies are to be printed 0: otherwise
23–24		blank
25–80	IDENT	Name (alphanumeric identification) of the problem

Format Card(s)

The Format Cards must provide for reading the c elements of a row at one time, using X, F, or E fields.

Input Data Cards

The Data Cards shall be punched in accordance with the format given on the Format Cards. The c elements of a row are read at one time from a set of cards. Each row starts on a new set of cards. Each data element is assumed to be integral valued. Any number of sets of Data Cards can be processed as one problem if r and c are the same for all sets and the data elements follow the same format for all sets.

continued

Finish Card

Columns	*Contents*
1– 6	FINISH Used after last problem (see next section)

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

Program Control Cards for problem 1
   Problem Card for problem 1
   Format Card(s) for problem 1
Data Cards for problem 1
Program Control Cards for problem 2
Data Cards for problem 2

   ⋮

Program Control cards for last problem
Data Cards for last problem
Finish Card

SAMPLE INPUT

```
PROBLEM  1  3  4  1  1  1  UMST520 SAMPLE PROBLEM
(4F2.0)
 6 4 3 2
 4 2 2 1
 3 1 3 1
FINISH
```

SAMPLE OUTPUT

```
UMST520  CHI-SQUARE UMST520 SAMPLE PROBLEM


   1 PROBLEM   3 ROWS   4 COLUMNS   1 FORMAT CARDS
      MARGINAL SUMS  1 =  YES.  EXPECTED FREQUENCIES  1 =  YES.

THIS PROBLEM REQUIRES A FIELD LENGTH OF 015606B

      FORMAT = (4F2.0)



PROBLEM 1     CHI-SQUARE =  1.2324E+00    DEGREES OF FREEDOM =  6

    ROW TOTALS             I =  3
             15          9        8

    COLUMN TOTALS              J =  4
             13          7        8        4

    GRANT TOTAL =             32


EXPECTED FREQUENCIES
ROW 1      6.094      3.281      3.750      1.875
ROW 2      3.656      1.969      2.250      1.125
ROW 3      3.250      1.750      2.000      1.000
```

COST ESTIMATE

For the job listed on Sample Input, the central processor time used was 0.100 seconds. At the current rate for the University of Minnesota ($0.20/sec.), the computer time cost $0.02 plus a small charge for output supplies.

Charge to user = computer cost + postage + network overhead
= $0.04 + postage + network overhead

CONTENTS—UMST520

pages

1– 2 Identification & Abstract
3– 4 User Instructions
5– 6 I/O
7 Cost—Contents

DESCRIPTIVE TITLE	Missing Data Correlation
CALLING NAME	UMST530
INSTALLATION NAME	University Computer Center University of Minnesota
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC COMPASS
COMPUTER	CDC 6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Technical Representative, Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Mn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

UMST530 computes Pearsonian product-moment correlations for all possible pairwise combinations of up to 130 variables. For this statistic and for the matrix of covariances, missing data for one or both variables of a pair exclude that observation from the analysis for that pair only; other pairs with non-missing data are retained. A count of non-missing observations for each set of pairs is kept and output. For each individual variable a count is also kept and output of the non-missing observations. For those observations a mean, variance and standard deviation are output.

Formulae

Let X_{ni} denote the nth observation of variable i where n = 1, 2, ..., NOBS and i = 1, 2, ..., NVAR, i.e. a total of NVAR variables with NOBS observations. Let $A_{nij} = 0$ if either X_{ni} or X_{nj} is missing and $A_{nij} = 1$ otherwise.

Count: $$N_{ij} = \sum_{n=1}^{NOBS} A_{nij}$$

continued

Mean: $\overline{X}_{ij} = \frac{1}{N_{ij}} \sum_{n=1}^{NOBS} A_{nij}X_{ni}$

Variance: $v_{ij} = \frac{1}{N_{ij}} \sum_{n=1}^{NOBS} A_{nij}(X_{ni} - \overline{X}_{ni})^2$

Covariance: $c_{ij} = \frac{1}{N_{ij}} \sum_{n=1}^{NOBS} A_{nij}(X_{ni} - \overline{X}_{ij})(X_{nj} - \overline{X}_{ji})$

Correlation: $r_{ij} = c_{ij} / (v_{ij}v_{ji})^{1/2}$

The user must note that $N_{ij} = 0$ is possible and that output results will be zero for that i and j, and mean nothing. Also, if $v_{ij} = 0$ (or $v_{ji} = 0$), the correlation will be zero and mean nothing.

USER INSTRUCTIONS

System Control Cards (prepared by Univ. of Minn. personnel)

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8–10	NVAR	Number of variables (does not include the counting variable) (NVAR $\geq$ 2)
11	MODE	0: Blanks on data fields are read as zeros. Missing Data Identification Cards are required. 1: All-blank and all-non-numeric fields are considered as missing. No Missing Data Identification Cards necessary. Warning: non-numeric characters are eliminated. Thus 4.0 and 0.4 become 40 and 4 respectively.
12	NFC	Number of Format Cards provided. If blank or zero, NFC=1 will be assumed. Hence there must be at least one Format Card.
13–15	NC	Indicates which variable position is used for counting observations (or individuals). If NC is blank or zero, NC=1 is assumed (see Format Cards and End of Data Cards below). A minus one in this variable position on the input data terminates input data reading. ($0 \leq NC \leq NVAR+1$)
21–80	NAME	Alphanumeric identification of the problem.

Format Card(s)

The Format Card(s) must provide for reading all NVAR variables (plus the counting variable) for one observation (or individual) at one time. If MODE = 1, X or A fields are to be used with an I field for counting observations. If MODE = 0, F, E, or X

continued

fields are necessary with the I field for counting observations. It is important that *exactly* NC-1 F or E fields (or A fields) occur prior to the I field in order that the observation-counting functions properly.

Missing Data Identification Cards

(Necessary if Col. 11 of the Problem Card is zero or blank.) A value for each variable is punched on this card (or cards) in accordance with the Format Card(s). If the value of a variable on a data card is the same as the value for that variable on the Missing Data Identification Card, the variable on the data card will be identified as missing. Each data card set is matched with the Missing Data Identification Card set to identify missing items.

Input Data

The data cards (or coded records on tape or on the disk) must be punched in accordance with the format given on the Format Card(s). Let the Input Data consist of NVAR variables with NOBS observations (or individuals). Hence an entry of the data is X_{ij} where $i = 1, 2, \ldots,$ NOBS and $j = 1, 2, \ldots,$ NVAR. The Format Card shall provide for reading in, from a set of cards (or a set of coded records from a tape or disk file), the values $X_{i1}, X_{i2}, X_{i3}, \ldots, X_{iNVAR}$. The program will then read NOBS such sets of cards, and the End of Data Card(s) should be the NOBS+1 observation. In other words, data are read one observation (or individual) at a time.

End of Data Card(s)

As explained above, the End of Data Card(s) consist(s) of the NOBS + 1 observation. The card(s) has (have) a -1 punched in the columns specified by the Format Card(s) (the minus sign is the 11-level punch). If more than one card per observation exists, it is necessary to have the same number of End of Data Cards, as the entire format is read prior to the test for the minus one.

Finish Card

The Finish Card has FINISH in Cols. 1–6, and terminates the search for the next Problem Card.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

continued

Problem Card for problem 1
Format Cards for problem 1
Missing Data Identification Cards (if needed) for problem 1
Data cards for problem 1
End of Data Card(s) for problem 1
Program Control Cards for problem 2
Data cards for problem 2
  Repeat as desired for problems 3, 4, etc.
Finish Card

Time Allocation

The time used by this program is given approximately by the formula:

$$T = \frac{(\text{NOBS})(\text{NVAR})^{1.88}}{1.5 \times 10^5} \text{ seconds}$$

543

9

SAMPLE INPUT

```
PROBLEM  301  4      MODE 0 UMST530  SAMPLE PROBLEM
(3(2X,F4.0),I2)

  9999  9999  0000
   300  2100   -40
  -400  9999  -100
  4200   500   600
  2100  0000   100
  9999   800  0000
                 -1
PROBLEM  311  4      MODE 1 UMST530  SAMPLE PROBLEM
(3(2X,A4),I2)
   300  2100   -40
         800
  4200   500   600
  -400        -100
  2100  0000   100
                 -1
FINISH
```

SAMPLE OUTPUT

```
UMST530       MISSING DATA CORRELATION              MODE 0 UMST530 SAMPLE PROBLEM
TOTAL VARIABLES =   3         MODE=   0         NUMBER FORMAT CARDS=  1         COUNT VARIABLE =     4


       FORMAT=  (3(2X,F4.0),I2)
THIS PROBLEM REQUIRES A FIELD LENGTH OF 020000


NUMBER OF OBSERVATIONS =          6

                         VAR    COUNT          MEAN                   STD DEV

                          1       6        4366.333333            4240.102226
                          2       5        4679.600000            4374.453386
                          3       4        140.0000000            275.3179980



     MATRIX OF COUNTS

 1           6

 2           5         5

 3           4         3         4



     COVARIANCE MATRIX

 1              17978466.89

 2             -1211879.960           19153344.24

 3              472500.0000          -905248.8889           75800.00000



     CORRELATION MATRIX

 1        1.000000

 2       -.061397    1.000000

 3        .963533    -.688166    1.000000
```

continued

```
UMST530        MISSING DATA CORRELATION                MODE 1 UMST530 SAMPLE PROBLEM
TOTAL VARIABLES =   3          MODE=   1          NUMBER FORMAT CARDS=  1          COUNT VARIABLE =     4


         FORMAT=  (3(2X,A4),I2)
THIS PROBLEM REQUIRES A FIELD LENGTH OF 020000


NUMBER OF OBSERVATIONS =          5

                          VAR     COUNT            MEAN                          STD DEV

                           1        4          1550.000000                     1781.151313
                           2        4          850.0000000                     776.2087348
                           3        4          140.0000000                     275.3179980



     MATRIX OF COUNTS

  1          4

  2          3        4

  3          4        3        4



     COVARIANCE MATRIX

  1              3172500.000

  2              -996666.6667              602500.0000

  3               472500.0000             -118666.6667              75800.00000



     CORRELATION MATRIX

  1         1.000000

  2         -.698209    1.000000

  3          .963533    -.482285    1.000000
```

COST ESTIMATE

For the job listed on the Sample Input, the central processor time used was 0.168 seconds. At the current rate for the University of Minnesota ($0.20/sec.), the computer time cost $0.03 plus a small charge for output supplies.

Charge to user = computer cost + postage + network overhead
= $0.27 + postage + network overhead

CONTENTS—UMST530

pages	
1– 2	Identification & Abstract
3– 5	User Instructions
7– 9	I/O
11	Cost—Contents

DESCRIPTIVE TITLE	Nonparametric (Rank Order) Statistics
CALLING NAME	UMST540
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	David Braddock Museum of Natural History Micheal Frisch University Computer Center
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Technical Representative, Center for Urban and Regional Affairs, University of Minnesota, 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program is designed to rank input data, then calculate and print selected rank-order statistics. Any completely non-numeric (including blank) input field is considered to be missing data and calculations involving that field are not made. In making this check all non-numerics (except a leading minus) are stripped from the number, which will produce inaccurate results for the user who has non-aligned punched decimals, e.g. 0.8953 and 59.0 will be accepted as 8953 and 590, thus reversing the order of the numbers.

The problem size limitation is a function of the number of variables and the number of observations: 2 variables and 5000 observations or 100 variables and 100 observations both approach this limit. Statistics include:

1. Kruskal-Wallis One-Way Analysis of Variance both corrected for ties and uncorrected for *K* independent samples. Where appropriate, the chi-square probability is also printed. The ranking in this option is performed over all elements of data except for missing data.

continued

2. Spearman Rank Correlation between all pairs of variables over all observations where both variables represent non-missing data. The t statistic and the Student's t probability may also be output. R_s is fully corrected for ties.

3. Kendall Rank Correlation. All statements made in 2 above apply here except that the statistics printed are tau, Z, S and the normal distribution probability.

4. Kendall Coefficient of Concordance with its associated chi-square statistic and probability. As with 2 and 3 above, the ranking is performed within each separate variable, but here a single missing variable will delete an entire observation from the analysis.

Computation and Theory

The theory and computations of this program are to be found in Ref. 1.

Kruskal-Wallis One-Way Analysis of Variance

The statistic H is defined by the formula:

$$H(\text{uncorrected for ties}) = \frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j} - 3(N+1)$$

where k = the number of variables
n_j = the number of observations for the jth variable (excluding missing observations)
$N = \sum_{j=1}^{k} n_j$, the total number of observations (excluding missing observations)
R_j = the sum of the ranks of observation for the jth variable

H corrected for ties is obtained by dividing the uncorrected H by

$$1 - \frac{\sum_{m=1}^{M} t_m^3 - t_m}{N^3 - N}$$

where t_m is the number of tied values in a tied group of observations and M is the number of tied groups.

Test of significance of H

Siegel[1] suggests that if not all $n_j \leq 5$, or $k \neq 3$, then a chi-

continued

square table may be used to test the significance of H(d.f. = k–1) (Table C in Siegel). In this case, the chi-square probability is printed by UMST540. Siegel's table 0 contains exact probabilities which can be used to test the significance of H when all $n_j \le 5$ and $k = 3$.

Spearman Rank Correlation Coefficient for two variables, x and y

$$r_s = \frac{\Sigma x^2 + \Sigma y^2 - \Sigma d_i^2}{2\sqrt{\Sigma x^2 \; \Sigma y^2}} \quad \text{this formula corrects for ties)}$$

where

$$\Sigma x^2 = \frac{N^3 - N}{12} \quad T_x$$

where N is the number of observations, (missing data in one variable cause deletion of corresponding data in the other variable), and

$$T_x = \sum_{m=1}^{M_x} \frac{t_m^3 - t_m}{12} \quad \text{for the x variable}$$

t_m is the number of tied values in a tied group of observations and M_x is the number of tied groups in the x variable;

Σy^2 is the y-variable analog of x^2 ;

$\sum_{i=1}^{N} d_i^2$ is the numeric difference between rank on the x-variable and rank on the y-variable for observation i.

Test of significance of r_s

For large N (> 10, according to Siegel), the following statistic is distributed as Student's t with d.f. = N–2:

$$t = r_s \sqrt{\frac{N-2}{1-r_s^2}}$$

and a table of Student's t can be used to test the significance of this statistic. In this case, the Student's t probability

continued

is printed by UMST540. For smaller N, Siegel suggests that r_s be tested against an exact probability table. Table P in Siegel lists exact probabilities for $N \leq 30$.

Kendall Rank Correlation Coefficient (tau) for two variables, x and y

$$\text{Tau} = \frac{S}{\sqrt{1/2\ N(N-1) - T_x}\ \sqrt{1/2\ N(N-1) - T_y}}$$

where N is defined as for r_s above,

$$T_x = 1/2 \sum_{m=1}^{M_x} t_m\,(t_m - 1) \quad \text{for variable } x\ ,$$

T_y is the y-analog of T_x,

t_m, M_x and M_y are defined as for r_s above.

S is derived as follows: Let x be the vector of ranks of A, y be the vector of ranks of B. Place the elements of x in ascending order. Place the elements of y associated with those of x in the same order.

Then

$$\begin{aligned} S = {} & \text{SIGN}\,[y_2-y_1] + \text{SIGN}\,[y_3-y_1] + \ldots + \text{SIGN}\,[y_n-y_1] \\ & + \text{SIGN}\,[y_3-y_2] + \text{SIGN}\,[y_4-y_2] + \ldots + \text{SIGN}\,[y_n-y_2] \\ & + \ldots + \ldots + \text{SIGN}\,[y_n-y_{n-1}] \end{aligned}$$

where SIGN $[y_i - y_j]$ = 1 if the expression in brackets is positive

=-1 if the expression in brackets is negative

= 0 if the expression in brackets is zero or if $x_i = x_j$

Test of Significance of Tau

For large N (Siegel suggests $N \geq 8$), Tau is distributed approximately as the normal distribution, and its significance may thus

continued

be tested with tables of the normal distribution, using

$$Z = \frac{\text{Tau}}{\frac{2(2N + 5)}{9N(N - 1)}} \quad \text{which has zero mean and unit variance}$$

In this case, the normal distribution probability is printed by UMST540. For smaller N ($\leq$ 10) Siegel's table Q may be used. This tests not the statistic Tau, but S, upon which Tau is based.

Kendall Coefficient of Concordance (W)

$$W = \frac{S}{\frac{1}{12} K^2(N^3 - N) - KT}$$

where N is the number of observations
K is the number of variables

$$S = \sum_{j=1}^{N} \left(R_j - \frac{\sum_{i=1}^{N} R_i}{N}\right)^2$$

where R_j is the sum of the ranks of observations for the jth variable

$$T = \sum_{j=1}^{K} T_j \quad \text{and} \quad T_j = \sum_{m=1}^{M_j} \frac{t_m^3 - t_m}{12} ,$$

where t_m is the number of tied values in a tied group of observations and M_j is the number of tied groups for the jth variable

Test of Significance of W

For samples greater than 7 the statistic K(N–1)W is distributed as χ^2 with degrees of freedom equal to N–1. In this case, the chi-square probability is printed by UMST540. For small samples ($\leq$ 7) Siegel's table R may be used.

REFERENCES

1. Siegel, S., *Non-Parametric Statistics* (New York: McGraw-Hill Book Co., Inc., 1956).

2. Kendall, M.G., *Rank Correlation Methods* (London: Charles Griffin & Co., Ltd., 1948).

USER INSTRUCTIONS

System Control Card

The System Control Card will be prepared by Univ. of Minnesota personnel.

Program Control Cards

Problem Control Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8– 9	NVAR	Number of variables (does not include optional counting variable) (NVAR $\geq$ 1)
10–13	NOBS	Number of observations. For K-W ANOVA, use maximum number of elements for a variable. If NOBS is zero or blank, the counting option is used (computer counts observations).
16	NFC	Number of Format Cards $0 \leq NFC \leq 9$ (if NFC = 0 or blank, 1 Format Card is assumed)
19	KRUSYES	1: Kruskal-Wallis is desired 0: otherwise
22	IPRNRNK	1: ranking printout option is desired 0: otherwise *Warning:* Large amounts of output are generated if this option is used.
25	ISPRYES	1: Spearman correlation coefficients are desired 0: otherwise
28	KENDYES	1: Kendall correlation coefficients are desired 0: otherwise
40	KONKORD	1: Kendall coefficient concordance is desired 0: otherwise

continued

Columns	*Parameter*	*Contents*
49–50	NC	The variable number in the format corresponding to where a –1 appears on the End of Data Cards. For example, if the –1 appears in Col. 9 – 10 and two data variables precede that field and three others follow it, the format could be (1X, F2.0, 1X, F2.0, 2X, I2, 3(1X, F2.0)) and NC is 3. The total number of observations is printed out at the end. $1 \le NC \le NVAR + 1$
51–80	NAME	Alphanumeric identification of the problem.

Format Card(s)

The Format Card(s) must provide for reading all NVAR variables for one observation at one time. The data elements may *only* be read by using *A* fields except that if the counting option is used, an I-field must be used where the –1 on the End of Data Card appears.

Input Data Cards

The Data Cards must be punched in accordance with the format given on the Format Card(s). Let X_{ni} denote the nth observation of variable i where n = 1,2, ..., NOBS and i = 1, 2, ..., NVAR. The Format Card must provide for reading in, from a set of cards, the values X_{n1}, X_{n2}, ..., $X_{n,NVAR}$. The program will then read in NOBS such sets of cards. The values within each variable are independent of values in another variable in the Kruskal-Wallis One Way Analysis of Variance. For this case assume a correspondence between values in different variables. Thus, if variable A has more data elements than variable B, assume missing data in variable B correspond with some of the observations in variable A. NOBS in this case is the maximum number of observations for any variable.

If decimal values are used as input data, it is important that the same number of places behind the decimal be used in all values. Example: the number 73. and .475 will be ranked as though the actual values were 73 and 475 and the resulting ranking will be wrong. Correct ranking will occur if the numbers are 73.000 and .475. Use extreme caution if input data are in decimal form.

continued

End of Data Card

An End of Data Card observation is assumed if NOBS=0 or if NOBS is blank. NC should be the variable number on the format corresponding to where the minus sign is punched on the End of Data Card. As many blank cards as are needed to complete an observation must be included with the End of Data Card.

Finish Card

The word FINISH is punched in Col. 1 – 6.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

```
Problem Control Card for problem 1                    ┐
Format Card(s) for problem 1                          │ Repeat as desired
Data Cards for problem 1                              │ for problem 2, 3,
End of Data Card for problem 1 (if necessary)         ┘ etc.
Finish Card
```

SAMPLE INPUT

```
PROBLEM 8         1  1                           1K-W ANOVA SEE SIEGEL P 190-2
(I2,8A4)
  20  35  33  32  26  31  26  25
  28  28  36  33  26  29  22  24
  33  32  26  32  29  31  22  30
  32  35  31  29  20  25  25  15
  44  23  32  33  20      12
  36  24  33  25  21      12
  19  20  29  26
  33  16  34  28
  28      32
  11      32
-1
PROBLEM 2      1  1  1  1  1                     3SPEAR+KEND RANK-SEE SIEGEL 205
(2A4,I2)
  82  42
  98  46
  87  39
  40  37
 116  65
 113  88
 111  86
  83  56
  85  62
 126  92
 106  54
 117  81
        -1
PROBLEM 3      1     1     1           1         1TEST KONKORD SCALING P234 SIEG
(I2,3(A2,2X))
   2   5   4
   9   2   2
   4   5   9
   9   9   9
   6   9   9
  15  16   9
  12  18  16
  18  13  16
  15  20  16
  20  13  20
-1
FINISH
```

SAMPLE OUTPUT

```
UMST540  NON-PARAMETRIC RANK ORDER STATISTICS   K-W ANOVA SEE SIEGEL P 190-2

       NUMBER OF VARIABLES =     8
       MAXIMUM NUMBER OF OBSERVATIONS IN ONE VARIABLE =    -0
       NUMBER OF FORMAT CARDS = -0
            KRUSKAL-WALLIS ONE-WAY ANOVA       1 = YES
            RANKING PRINT-OUT                  1 = YES
            SPEARMAN RANK CORRELATION         -0 = NO
            KENDALL RANK CORRELATION          -0 = NO
            KENDALL COEFFICIENT OF CONCORD    -0 = NO
       VARIABLE NUMBER OF COUNT VARIABLE =     1

       FORMAT = (I2,8A4)

          10 OBSERVATIONS READ

       THIS PROBLEM REQUIRES A FIELD LENGTH OF 031700B

VARIABLE  MEDIAN RANK  NO. OF OBS.  INDIVIDUAL RANKS   (ZEROS ARE FOR MISSING DATA)

   1          5.25        10     3.00     4.50     7.50     6.00    10.00     9.00     2.00     7.50     4.50     1.00

   2          4.50         8     7.50     5.00     6.00     7.50     3.00     4.00     2.00     1.00     0.00     0.00

   3          5.00        10     7.50    10.00     1.00     3.00     5.00     7.50     2.00     9.00     5.00     5.00

   4          4.75         8     5.50     7.50     5.50     4.00     7.50     1.00     2.00     3.00     0.00     0.00

   5          3.75         6     4.50     4.50     6.00     1.50     1.50     3.00     0.00     0.00     0.00     0.00

   6          2.75         4     3.50     2.00     3.50     1.00     0.00     0.00     0.00     0.00     0.00     0.00

   7          3.50         6     6.00     3.50     3.50     5.00     1.50     1.50     0.00     0.00     0.00     0.00

   8          2.50         4     3.00     2.00     4.00     1.00     0.00     0.00     0.00     0.00     0.00     0.00

           MEDIAN RANK  NO. OF OBS.  INDIVIDUAL RANKS   (ZEROS ARE FOR MISSING DATA)

             34.25        10     8.50    27.50    47.50    41.00    56.00    54.50     6.00    47.50    27.50     1.00

             21.50         8    52.50    27.50    41.00    52.50    14.00    15.50     8.50     5.00     0.00     0.00

             41.00        10    47.50    54.50    23.00    36.00    41.00    47.50    31.50    51.00    41.00    41.00

             36.25         8    41.00    47.50    41.00    31.50    47.50    18.50    23.00    27.50     0.00     0.00

             17.00         6    23.00    23.00    31.50     8.50     8.50    11.00     0.00     0.00     0.00     0.00

             33.75         4    36.00    31.50    36.00    18.50     0.00     0.00     0.00     0.00     0.00     0.00
```

continued

```
          12.50         6    23.00    12.50    12.50    18.50     2.50     2.50     0.00     0.00     0.00     0.00


          17.00         4    18.50    15.50    34.00     4.00     0.00     0.00     0.00     0.00     0.00     0.00

        KRUSKAL-WALLIS H STATISTIC =          18.4639   CHI-SQUARE PROBABILITY =          .9900
        H CORRECTED FOR TIES       =          18.5654   CHI-SQUARE PROBABILITY =          .9904
        DEGREES OF FREEDOM         =                7



UMST540  NON-PARAMETRIC RANK ORDER STATISTICS   SPEAR+KEND RANK-SEE SIEGEL 205

        NUMBER OF VARIABLES =     2
        MAXIMUM NUMBER OF OBSERVATIONS IN ONE VARIABLE =   -0
        NUMBER OF FORMAT CARDS =  1
             KRUSKAL-WALLIS ONE-WAY ANOVA       1 = YES
             RANKING PRINT-OUT                  1 = YES
             SPEARMAN RANK CORRELATION          1 = YES
             KENDALL RANK CORRELATION           1 = YES
             KENDALL COEFFICIENT OF CONCORD    -0 = NO
        VARIABLE NUMBER OF COUNT VARIABLE =     3

        FORMAT = (2A4,I2)

           12 OBSERVATIONS READ

        THIS PROBLEM REQUIRES A FIELD LENGTH OF 031700B

VARIABLE  MEDIAN RANK  NO. OF OBS.  INDIVIDUAL RANKS   (ZEROS ARE FOR MISSING DATA)

   1           6.50        12    2.00     6.00     5.00     1.00    10.00     9.00     8.00     3.00     4.00    12.00
                                 7.00    11.00

   2           6.50        12    3.00     4.00     2.00     1.00     8.00    11.00    10.00     6.00     7.00    12.00
                                 5.00     9.00

          MEDIAN RANK  NO. OF OBS.  INDIVIDUAL RANKS   (ZEROS ARE FOR MISSING DATA)

              18.50        12   11.00    18.00    15.00     3.00    22.00    21.00    20.00    12.00    13.00    24.00
                                19.00    23.00

               7.50        12    4.00     5.00     2.00     1.00     9.00    16.00    14.00     7.00     8.00    17.00
                                 6.00    10.00

        KRUSKAL-WALLIS H STATISTIC =           8.6700   CHI-SQUARE PROBABILITY =          .9968
        H CORRECTED FOR TIES       =           8.6700   CHI-SQUARE PROBABILITY =          .9968
        DEGREES OF FREEDOM         =                1
  FIRST      MEDIAN    SECOND     MEDIAN       NUMBER            SPEARMAN RANK CORRELATION          KENDALL RANK CORRELATION
 VARIABLE     RANK     VARIABLE    RANK        OF OBS.          RS          T    STUDENT T PROB    TAU         Z      NORMAL PROB
    1         6.50        2        6.50          12         8.18E-01    4.50E+00     .99845     6.67E-01   3.02E+00     .99872

 VARIABLE       1          SORTED RANKS      1.00     2.00     3.00     4.00     5.00     6.00     7.00     8.00     9.00    10.00
                                            11.00    12.00

 VARIABLE       2    CORRESPONDING RANKS     1.00     3.00     6.00     7.00     2.00     4.00     5.00    10.00    11.00     8.00
                                             9.00    12.00
```

continued

```
UMST540  NON-PARAMETRIC RANK ORDER STATISTICS   TEST KONKORD SCALING P234 SIFG

         NUMBER OF VARIABLES =     3
         MAXIMUM NUMBER OF OBSERVATIONS IN ONE VARIABLE =    -0
         NUMBER OF FORMAT CARDS =  1
              KRUSKAL-WALLIS ONE-WAY ANOVA       -0 = NO
              RANKING PRINT-OUT                   1 = YES
              SPEARMAN RANK CORRELATION          -0 = NO
              KENDALL RANK CORRELATION            1 = YES
              KENDALL COEFFICIENT OF CONCORD      1 = YES
         VARIABLE NUMBER OF COUNT VARIABLE =      1

         FORMAT = (I2,3(A2,2X))

            10 OBSERVATIONS READ

         THIS PROBLEM REQUIRES A FIELD LENGTH OF 031700B

VARIABLE  MEDIAN RANK  NO. OF OBS.  INDIVIDUAL RANKS   (ZEROS ARE FOR MISSING DATA)

    1          5.25        10     1.00     4.50     2.00     4.50     3.00     7.50     6.00     9.00     7.50    10.00

    2          5.50        10     2.50     1.00     2.50     4.50     4.50     8.00     9.00     6.50    10.00     6.50

    3          4.50        10     2.00     1.00     4.50     4.50     4.50     4.50     8.00     8.00     8.00    10.00

 FIRST     MEDIAN    SECOND     MEDIAN        NUMBER              KENDALL RANK CORRELATION
VARIABLE    RANK    VARIABLE    RANK         OF OBS.              TAU          Z       NORMAL PROB
    1       5.25        2       5.50            10             4.71E-01    1.89E+00      .97090

VARIABLE      1         SORTED RANKS      1.00     2.00     3.00     4.50     4.50     6.00     7.50     7.50     9.00    10.00

VARIABLE      2   CORRESPONDING RANKS     2.50     2.50     4.50     4.50     1.00     9.00     8.00    10.00     6.50     6.50

 FIRST     MEDIAN    SECOND     MEDIAN        NUMBER              KENDALL RANK CORRELATION
VARIABLE    RANK    VARIABLE    RANK         OF OBS.              TAU          Z       NORMAL PROB
    1       5.25        3       4.50            10             6.61E-01    2.66E+00      .99609

VARIABLE      1         SORTED RANKS      1.00     2.00     3.00     4.50     4.50     6.00     7.50     7.50     9.00    10.00

VARIABLE      3   CORRESPONDING RANKS     2.00     4.50     4.50     4.50     1.00     8.00     4.50     8.00     8.00    10.00

 FIRST     MEDIAN    SECOND     MEDIAN        NUMBER              KENDALL RANK CORRELATION
VARIABLE    RANK    VARIABLE    RANK         OF OBS.              TAU          Z       NORMAL PROB
    2       5.50        3       4.50            10             6.69E-01    2.69E+00      .99644

VARIABLE      2         SORTED RANKS      1.00     2.50     2.50     4.50     4.50     6.50     6.50     8.00     9.00    10.00

VARIABLE      3   CORRESPONDING RANKS     1.00     2.00     4.50     4.50     4.50    10.00     8.00     4.50     8.00     8.00

VARIABLE  MEDIAN RANK  NO. OF OBS.  INDIVIDUAL RANKS   (ZEROS ARE FOR MISSING DATA)

    1          5.25        10     1.00     4.50     2.00     4.50     3.00     7.50     6.00     9.00     7.50    10.00
```

continued

```
   2          5.50      10     2.50      1.00      2.50      4.50      4.50      8.00      9.00      6.50     10.00      6.50

   3          4.50      10     2.00      1.00      4.50      4.50      4.50      4.50      8.00      8.00      8.00     10.00
KENDALL COEFFICIENT OF CONCORDANCE IS          .8277   CHI-SQUARE STATISTIC   =        22.3487
WITH DEGREES OF FREEDOM                       =     9   CHI-SQUARE PROBABILITY =          .9922
```

COST ESTIMATE

For the job listed on the Sample Input, the central processor time used was 0.498 seconds. At the current rate for the University of Minnesota ($0.20/sec.), the computer time cost $0.10 plus a small charge for output supplies.

Charge to user = computer time + postage + network overhead
= $0.22 + postage + network overhead

CONTENTS—UMST540

pages

1– 5	Identification & Abstract
7– 9	User Instructions
11–15	I/O
17	Cost—Contents

DESCRIPTIVE TITLE	FORTRAN Program for Computer Based Serials Holdings Management
CALLING NAMES	UPDATE, HOLD, PUBLSH, LANSUB
INSTALLATION NAME	Indiana University-Purdue University at Indianapolis Research Computation Center
AUTHOR(S) AND AFFILIATION(S)	Mrs. Alma Connell School of Medicine Library Indiana University Mrs. J. Mueller School of Medicine Library Indiana University Mrs. Judy Silence Research Computation Center, IUPUI
LANGUAGE	IBM FORTRAN IV (IBSYS)
COMPUTER	IBM 7040
PROGRAM AVAILABILITY	Deck and listing presently available
CONTACT	Dr. David A. Neal, EIN Tech. Rep., Research Comp. Ctr., Indiana Univ.–Purdue Univ. at Indianapolis, 1100 West Michigan St., Indianapolis, Ind. 46202 Tel.: (317) 639-7813

FUNCTIONAL ABSTRACT

This system is a long range project concerning the holdings of serials in the School of Medicine Library. Three overall applications are (1) updating the file to maintain a current holding list, (2) publication of selected lists using search techniques, and (3) annual publication of complete holdings list.

Several programs have been prepared for these applications. Detailed descriptions are given in Ref. 1. A general description of each program follows.

System Update Program 1

This program produces a readable list for visual checking of every item on the cards. When a set of cards has been keypunched,

continued

they are run with the System Update Program 1 using the card reader as the input unit. The output is checked and the cards are corrected. The cards are then added to the master file with the UPDATE (number 2) program.

System Update Program 2

This program builds a master tape which may be used as data for the programs that follow. UPDATE will also be used for maintaining the current file, i.e., inserting new cards as they are punched and replacing cards when new information is needed.

HOLD

HOLD can be run with the master file to produce a listing of the information on each card on that file. HOLD also provides a summary table giving the number of titles for specified categories.

PUBLSH

PUBLSH provides thirteen options for listing holdings for publication. These are (1) full list with holdings, (2 and 3) nursing list with or without holdings, (4 and 5) currently received list with or without holdings, (6 and 7) Index Medicus list with or without holdings, (8 and 9) Indexes, Abstracts and Bibliographies list with or without holdings, (10 and 11) International Nursing Index with or without holdings, and (12 and 13) microform with or without holdings. Multiple copies of output from this program are available.

LANSUB

LANSUB provides four general options for title listings. These are (1) language only, (2) one subject, (3) two subjects listed according to "and" or "or" logic, and (4) a language and a subject cross-indexed.

Listings of complete holdings or of currently received serials only can be obtained under the four options above. On complete holdings lists, currently received titles are indicated by a plus (+) to the left of the title. Multiple copies of output from this program are available.

REFERENCE

1. Indiana University Medical Center Research Computation Center Library, *Program Description* (Indianapolis, Ind: Indiana Univ., 1966, Rev. 1968). Available from the EIN Office at the cost of reproduction and mailing.

USER INSTRUCTIONS—SYSTEM UPDATE PROGRAM 1

Input Cards

UPDATE 1 produces a list of the Data Cards for verification. The data may be input in any order. After the program has been run, the cards are sorted. A full data set for one serial includes Cards 1–4.

Card 1

Columns	*Contents*
1–64	First part of entry. Keypuncher is responsible for logical break in entry if title goes on two cards
65–72	blank
73–79	Entry number; punch leading zeroes, if any
80	1

Card 2

Columns	*Contents*
1–58	Continuation of entry. Program will not print this line if Cols. 1–6 are blank.
59–66	blank
67	X: this is a cross reference blank: not a cross reference
68	S: supplements included blank: no supplements
69	+: serial is currently received blank: not currently received
70–72	Number of volumes
73–79	Entry number; punch leading zeroes
80	2

Card 3

Columns	*Contents*
1–58	Holdings and cross reference title. If Col. 67, Card 2 = X, this should be a primary–reference title; otherwise (Col. 67 = blank), this should contain the holdings information.

continued

Columns	*Contents*
59–65	blank
66	M: serial is on microform blank: not on microform
67	A: serial is indexes, abstracts or bibliographies blank: not of these types
68	N: nursing reference blank: not a nursing reference
69	Language blank: English DA: Danish DU: Dutch FR: French GE: German HE: Hebrew IT: Italian JA: Japanese RU: Russian SP: Spanish PO: Portuguese SC: Serbo-Croatian SL: Slovenian AL: Albanian NO: Norwegian PL: Polish
71	*: serial is listed in Index Medicus blank: not listed in Index Medicus
72	W: serial is listed in Inter. Nurs. Index (I.N.I.) blank: serial is not listed in I.N.I.
73–79	Entry number; punch leading zeroes
80	3

Card 4

Columns	*Contents*
1–58	Continuation of holdings & reference. This line is not printed if Cols. 1–6 are blank.
59–64	blank
65–66	First subject category, if any
67–68	Second subject category, if any

continued

Columns	*Contents*
69–70	Third subject category, if any
71–72	Fourth subject category, if any
73–79	Entry number; punch leading zeroes
80	4

Subject Categories

Code	*Subject*	*Code*	*Subject*
BF	Psychology	WD	Deficiency diseases
H	Sociology	W2	Metabolic diseases
HD	Rehabilitation	W3	Allergy
L	Education	W4	Animal poisoning
Q	Science	W5	Plant poisoning
QA	Math	W6	Physical agents, diseases caused by
QC	Physics		
QD	Chemistry	W7	Aviation & space med.
QH	Biology	WE	Musculoskeletal system
QK	Botany	WF	Respiratory system
QL	Zoology	WG	Cardiovascular system
QP	Physiology (animal)	WH	Hemic & lymphatic systems
QS	Human anatomy	WI	Gastrointestinal system
QT	Physiology	WJ	Urogenital system
QU	Biochemistry	WK	Endocrine system
QV	Pharmacology	WL	Nervous system, sense organs
QW	Bacteriology	WM	Psychiatry
QX	Parasitology	WN	Radiology
QY	Clinical pathology	WO	Surgery, anesthesia
QZ	Pathology	WP	Gynecology
S	Agriculture	WQ	Obstetrics
SF	Veterinary medicine	WR	Dermatology
T	Human engineering	WS	Pediatrics
UH	Military medicine	WT	Geriatrics
W	Medical profession	WV	Otorhinolaryngology
WA	Public Health	WW	Ophthalmology
WB	Practice of medicine	WX	Hospitals
WC	Infectious diseases	WY	Nursing
		WZ	History of Medicine

Additional subject codes and groups of codes have been added. The only limitations on the use of additional subject codes is that they must be alphanumeric codes of no more two letters.

continued

000 0140
(a)

System Control Cards

System Control Cards should be punched as they appear below.

```
↓Col. 1        ↓Col. 16
$JOB
$EXECUTE       UPDATE
$RUN INPUT 10
$OUTPUT S.SU05
$OBTAIN
    [Unsorted Input Cards]
$ENDRUN
$IBSYS
 FILE,INPUT/7,BLOCKSIZE/150,REWIND
 FILE,OUTPUT,BLOCKSIZE/150,REWIND
 RECORD,LENGTH/15,FIELD/(78,8)
 SORT,FILE/7,SEQ/C,ORD/3,FIELD/(2)
 SYSTEM,INPUT/S.SU05,MERGE/(B,C),OUTPUT/S.SU05,DISK
 OPTION,NOCKPT,NOTAPE
 END
$IBSYS
$EXECUTE       UPDATE
$RUN NOCOPY  10
$ASSIGN S.SU05   OPEN
$OBTAIN PRINT
$ENDRUN
$IBSYS
```

USER INSTRUCTIONS—SYSTEM UPDATE PROGRAM 2

Input Cards

UPDATE 2 is used to change the master tape; adding, deleting, or correcting information. Cards correcting or adding information should follow the formats specified for UPDATE 1. Deletion Cards should be punched in the following format.

Columns	*Contents*
1– 7	$DELETE
9–16	Entry number + card number (Cols. 73–80) of the first card to be deleted (or if a single card).
18–25	Entry number + card number of the last card to be deleted; if two or more *consecutive* cards are deleted.

All of these cards (deletion, addition, correction) should be in ascending order by the numbers in Cols. 73-80.

System Control Cards

See HOLD.

USER INSTRUCTIONS—HOLD

System Control Cards

HOLD can produce a listing of the information on each card, as well as a summary table giving the number of titles for specified categories.

System Control Cards should be punched as they appear below. The next to last card is the HOLD Control Card.

```
↓Col. 1         ↓Col. 16

$JOB
$PAUSE          MOUNT 'OLD' MASTER ON 4, 'NEW' MASTER ON 3
$EXECUTE        UPDATE
$RUN UPDATE 10
$OUTPUT S.SU04    S.SU04
$ASSIGN S.SU05    OPEN
      [UPDATE 2 Input Cards (sorted)]
$ENDRUN
$IBSYS
$PAUSE          MOUNT 'NEW' MASTER ON 3
$IBJOB
$FILE           'FTC04.',U04
      [HOLD binary deck, provided by RCC personnel]
$ENTRY
04 (Punch date, starting in Col. 4; either with slashes or
    with month spelled out.)
$IBSYS
```

USER INSTRUCTIONS—PUBLSH

Control Cards

PUBLSH provides options for listing holdings for publication.

Card 1

Columns	*Contents*
1—24	The date, punched either with slashes or with the month spelled out

If the date is not wanted, leave this card blank.

Card 2

Punch as many cards as needed. Each card should have the following format.

Columns	*Contents*
1	List option blank: full list N: nursing list +: currently received list *: Index Medicus list A: indices, abstracts & bibliographies list W: I.N.I. list M: microform list
2— 4	YES: print holdings blank: do not print holdings (this option does not apply to the full list selection)
5— 6	04
7— 9	Printed stencil option YES: pause to change printer tape to print on stencils blank: do not pause; printing is not done on stencils

System Control Cards

System Control Cards should be punched exactly as they appear below.

```
↓Col. 1        ↓Col. 16

$JOB
$PAUSE         MOUNT 'NEW' MASTER ON 4
```

continued

```
↓Col. 1          ↓Col. 16

$IBJOB
$FILE            'FTC04.',U05
$FILE            'FTC10.',U05
    [PUBLSH binary deck; supplied by RCC personnel]
$ENTRY
    [PUBLSH Control Cards]
$IBSYS
$EXECUTE COPIES
$COPIES nnnn U04       (where nnnn = number of print copies
$IBSYS                  desired, right-justified)
```

USER INSTRUCTIONS—LANSUB

Input Cards

LANSUB provides options for output listings. Each selection consists of a set of Control and Title Cards

Control Card

Columns	*Contents*
1– 2	Language code; see UPDATE 1, Card 3
3– 4	First subject coce; see UPDATE 1, Card 4
5– 6	Second subject code
7	+: only currently received titles desired blank: otherwise
8– 9	04
10	Logic indicator if two subjects are requested 1: "or" logic 2: "and" logic

Title Cards

There must be two Title Cards, even if blank. Information on these cards will be printed at the top of the list. Use Cols. 1–72 only. Title will be centered on the page if centered on the cards.

System Control Cards

System Control Cards should be punched exactly as they appear below.

```
↓Col. 1         ↓Col. 16

$JOB
$PAUSE          MOUNT 'NEW' MASTER ON 4
$IBJOB
$FILE           'FTC04.',U05
$FILE           'FTC10.',U04
    [LANSUB binary deck; provided by IUPUI personnel]
$ENTRY
    [LANSUB Input Cards]
$IBSYS
$EXECUTE COPIES
$COPIES nnnnU04  (where nnnn is the number of printed copies
$IBSYS            desired, right-justified)
```

SAMPLE INPUT

```
$JOB          001002EDUCON     5 RET. TO J.SILENCE             00025001
ABT                                                     X      00025002
ABSTRACTS OF BIOANALYTIC TECHNOLOGY                       PO   00025003
                                                               00025004
ACM                                                            00050001
                                                        X      00050002
ASSOCIATION FOR COMPUTING MACHINERY                            00050003
                                                               00050004
AIBS BULLETIN                                                  00075001
                                                        X      00075002
BIOSCIENCE                                                     00075003
                                                               00075004
AMA AMERICAN JOURNAL OF DISEASES OF CHILDREN                   00100001
                                                        X      00100002
AMERICAN JOURNAL OF DISEASES OF CHILDREN                       00100003
                                                               00100004
AMA ARCHIVES ...                                               00125001
                                                        X      00125002
ARCHIVES ...                                                   00125003
                                                               00125004
AMA JOURNAL OF DISEASES OF CHILDREN                            00150001
                                                        X      00150002
AMERICAN JOURNAL OF DISEASES OF CHILDREN                       00150003
                                                               00150004
AMA NEWS                                                       00175001
                                                          +    00175002
LIBRARY RETAINS CURRENT YEAR ONLY                              00175003
                                                        W      00175004
AORN JOURNAL                                                   00200001
                                                          +00300200002
1-       1963-                                          N  *   00200003
                                                        WY     00200004
ASHA MONOGRAPHS                                                00225001
                                                          +00700225002
1-       1950-                                             *   00225003
                                                        WV     00225004
ASHA REPORTS                                                   00250001
                                                          +00100250002
1-       1965-                                                 00250003
                                                        WV     00250004
ABHANDLUNGEN AUS DEM GESAMTGEBIETE DER MEDIZIN                 00275001
                                                           00100275002
1924                                                      GE   00275003
                                                        WB     00275004
ABHANDLUNGEN AUS DER AUGENHEILKUNDE UND IHREN GRENZGEBIETEN    00300001
                                                        X      00300002
BIBLIOTHECA OPHTHALMOLOGICA                                    00300003
                                                               00300004
ABHANDLUNGEN AUS DER KINDERHEILKUNDE UND IHREN GRENZGEBIETEN   00325001
                                                        X      00325002
BIBLIOTHECA PAEDIATRICA                                        00325003
                                                               00325004
ABHANDLUNGEN AUS DER NEUROLOGIE. PSYCHIATRIE. PSYCHOLOGIE UND  00350001
IHREN GRENZGEBIETEN                                        00100350002
42       1927                                             GE   00350003
                                                        WLWM   00350004
ABSTRACTS OF BACTERIOLOGY                                      00375001
                                                           00900375002
1-9      1917-25                                        A      00375003
                                                        QW     00375004
ABSTRACTS OF BIOANALYTIC TECHNOLOGY                            00400001
                                                        S+00500400002
1-13     1953-65                                        A      00400003
                                                        QY     00400004
ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE            00425001
                                                        X      00425002
U.S. PUBLIC HEALTH SERVICE. DIV. OF VENEREAL DISEASE.          00425003
ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE            00425004
ABSTRACTS OF HOSPITAL MANAGEMENT STUDIES                       00450001
                                                          +00100450002
1-       1965-                                          A      00450003
                                                        WX     00450004
```

SAMPLE
SERIAL CARDS

continued

000 0140

```
ABSTRACTS OF HUMAN DEVELOPMENTAL BIOLOGY                                00475001
                                                                  X     00475002
EXCERPTA MEDICA. SECTION 21. DEVELOPMENTAL BIOLOGY AND                  00475003
TERATOLOGY                                                              00475004
ABSTRACTS OF JAPANESE MEDICINE                                          00500001
                                                                     00600500002
1-2      1960-62                                                  A     00500003
                                                                  WB    00500004
ABSTRACTS OF MYCOLOGY                                                   00525001
                                                                    +   00525002
(11-      1967-                                                   A     00525003
                                                                  QW    00525004
ABSTRACTS OF SOVIET MEDICINE                                            00550001
                                                                     00200550002
5      1961                                                       A     00550003
                                                                  WB    00550004
ABSTRACTS OF SOVIET MEDICINE. CANCER RESEARCH                           00575001
                                                                     00100575002
5-9      1953-56                                                  A     00575003
                                                                  QZ    00575004
ABSTRACTS OF SOVIET MEDICINE. PART A. BASIC MEDICAL SCIENCES            00600001
                                                                     00400600002
1-4      1957-60                                                  A     00600003
                                                                  W     00600004
ABSTRACTS OF SOVIET MEDICINE. PART B. CLINICAL MEDICINE                 00625001
                                                                     00600625002
1-4      1957-60                                                  A     00625003
                                                                  WB    00625004
ABSTRACTS OF VITAMIN LITERATURE                                         00650001
                                                                  X     00650002
VITAMIN ABSTRACTS                                                   PO  00650003
                                                                        00650004
ABSTRACTS OF WORLD MEDICINE                                             00675001
                                                                    +02100675002
1-      1947-                                                     A     00675003
                                                                  WB    00675004
ABSTRACTS OF WORLD SURGERY, OBSTETRICS AND GYNECOLOGY                   00700001
                                                                     00700700002
1-11      1947-52                                                 A     00700003
                                                                  WOWPWQ00700004
ACADEMIE DE CHIRURGIE, PARIS. MEMOIRES                                  00725001
                                                                    +   00725002
90-      1964-                                                      FR* 00725003
            •                              •                               •
            •                              •                               •
            •                              •                               •
                                                                    +00601100002
1-      1957-                                                         * 01100003
SUPPLEMENTS 1-      1959-                                         WO    01100004
ACTA ANATOMICA                                                          01125001
                                                                    +04401125002
1-      1945-                                                         * 01125003
SUPPLEMENTS 1-      1944-                                         QS    01125004
                                                                        99999991
                                                                        99999992
                                                                        99999993
                                                                        99999994
```

```
04 NOVEMBER 24.1970
```

Hold Test Data Card

```
FRWO   04  1
     PRINT FRENCH SURGERY SERIALS (TEST OUTPUT)

GEWL   04  1
     PRINT GERMAN - NERVOUS SYSTEM SERIALS (TEST OUTPUT)

FROHOT 04  2
     PRINT FRENCH BIOLOGY AND PHYSIOLOGY (TEST OUTPUT)

STOP
```

Lansub Test Data Cards

```
 NOVEMBER 24,1970
   04
NYES04
MYES04
M  04
STOP
```

Publsh Test Data Cards

000 0140

SAMPLE OUTPUT—UPDATE

```
500104SILENCE   10                    UPDATE SUMMARY                           12/18/70   PAGE  1

$RUN INPUT   10

$OUTPUT S.SU04   S.SU04
ABT                                                                       00025001         INSERTED
                                                                X         00025002         INSERTED
ABSTRACTS OF BIOANALYTIC TECHNOLOGY                                PO     00025003         INSERTED
                                                                          00025004         INSERTED
ACM                                                                       00050001         INSERTED
                                                                X         00050002         INSERTED
ASSOCIATION FOR COMPUTING MACHINERY                                       00050003         INSERTED
                                                                          00050004         INSERTED
AIBS BULLETIN                                                             00075001         INSERTED
                                                                X         00075002         INSERTED
BIOSCIENCE                                                                00075003         INSERTED
                                                                          00075004         INSERTED
AMA AMERICAN JOURNAL OF DISEASES OF CHILDREN                              00100001         INSERTED
                                                                X         00100002         INSERTED
AMERICAN JOURNAL OF DISEASES OF CHILDREN                                  00100003         INSERTED
                                                                          00100004         INSERTED
AMA ARCHIVES ...                                                          00125001         INSERTED
                                                                X         00125002         INSERTED
ARCHIVES ...                                                              00125003         INSERTED
                                                                          00125004         INSERTED
AMA JOURNAL OF DISEASES OF CHILDREN                                       00150001         INSERTED
                                                                X         00150002         INSERTED
AMERICAN JOURNAL OF DISEASES OF CHILDREN                                  00150003         INSERTED
                                                                          00150004         INSERTED
AMA NEWS                                                                  00175001         INSERTED
                                                                   +      00175002         INSERTED
LIBRARY RETAINS CURRENT YEAR ONLY                                         00175003         INSERTED
                                                                W         00175004         INSERTED
AORN JOURNAL                                                              00200001         INSERTED
                                                                      +00300200002         INSERTED
1-       1963-                                                   N      * 00200003         INSERTED
                                                                WY        00200004         INSERTED
ASHA MONOGRAPHS                                                           00225001         INSERTED
                                                                      +00700225002         INSERTED
1-       1950-                                                          * 00225003         INSERTED
                                                                WV        00225004         INSERTED
ASHA REPORTS                                                              00250001         INSERTED
                                                                      +00100250002         INSERTED
1-       1965-                                                            00250003         INSERTED
                                                                WV        00250004         INSERTED
ABHANDLUNGEN AUS DEM GESAMTGEBIETE DER MEDIZIN                            00275001         INSERTED
                                                                       00100275002         INSERTED
1924                                                              GE      00275003         INSERTED
                                                                WB        00275004         INSERTED
ABHANDLUNGEN AUS DER AUGENHEILKUNDE UND IHREN GRENZGEBIETEN               00300001         INSERTED
                                                                X         00300002         INSERTED
BIBLIOTHECA OPHTHALMOLOGICA                                               00300003         INSERTED
                                                                          00300004         INSERTED
ABHANDLUNGEN AUS DER KINDERHEILKUNDE UND IHREN GRENZGEBIETEN              00325001         INSERTED
                                                                X         00325002         INSERTED
BIBLIOTHECA PAEDIATRICA                                                   00325003         INSERTED
                                                                          00325004         INSERTED
ABHANDLUNGEN AUS DER NEUROLOGIE, PSYCHIATRIE, PSYCHOLOGIE UND             00350001         INSERTED
IHREN GRENZGEBIETEN                                                    00100350002         INSERTED
42       1927                                                    GE       00350003         INSERTED
                                                                WLWM      00350004         INSERTED
ABSTRACTS OF BACTERIOLOGY                                                 00375001         INSERTED
                                                                       00900375002         INSERTED
1-9      1917-25                                                A         00375003         INSERTED
                                                                QW        00375004         INSERTED
ABSTRACTS OF BIOANALYTIC TECHNOLOGY                                       00400001         INSERTED
                                                                     S+00500400002         INSERTED
1-13     1953-65                                                A         00400003         INSERTED
                                                                QY        00400004         INSERTED
ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE                       00425001         INSERTED
                                                                X         00425002         INSERTED
```

continued

000 0140
(b)

```
500104SILENCE   10                          UPDATE SUMMARY                                    12/18/70   PAGE  2

U.S. PUBLIC HEALTH SERVICE. DIV. OF VENEREAL DISEASE.                         00425003        INSERTED
ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE                           00425004        INSERTED
ABSTRACTS OF HOSPITAL MANAGEMENT STUDIES                                      00450001        INSERTED
                                                                           +00100450002       INSERTED
1-      1965-                                                          A      00450003        INSERTED
                                                                       WX     00450004        INSERTED
ABSTRACTS OF HUMAN DEVELOPMENTAL BIOLOGY                                      00475001        INSERTED
                                                                       X      00475002        INSERTED
EXCERPTA MEDICA. SECTION 21. DEVELOPMENTAL BIOLOGY AND                        00475003        INSERTED
TERATOLOGY                                                                    00475004        INSERTED
ABSTRACTS OF JAPANESE MEDICINE                                                00500001        INSERTED
                                                                            00600500002       INSERTED
1-2      1960-62                                                       A      00500003        INSERTED
                                                                       WB     00500004        INSERTED
ABSTRACTS OF MYCOLOGY                                                         00525001        INSERTED
                                                                         +    00525002        INSERTED
(1)-      1967-                                                        A      00525003        INSERTED
                                                                       QW     00525004        INSERTED
ABSTRACTS OF SOVIET MEDICINE                                                  00550001        INSERTED
                                                                            00200550002       INSERTED
5      1961                                                            A      00550003        INSERTED
                                                                       WB     00550004        INSERTED
                  .                           .                                  .
                  .                           .                                  .
                  .                           .                                  .
500104SILENCE   10                          UPDATE SUMMARY                                    12/18/70   PAGE  4

ACOUSTICAL SOCIETY OF AMERICA. JOURNAL                                        01025001        INSERTED
                                                                      S+00901025002           INSERTED
32-      1960-                                                              * 01025003        INSERTED
                                                                      WV      01025004        INSERTED
ACTA ALLERGOLOGICA                                                            01050001        INSERTED
                                                                         +01301050002         INSERTED
1-      1948-                                                               * 01050003        INSERTED
SUPPLEMENTS 1-7      1950-60                                          W3      01050004        INSERTED
ACTA ANAESTHESIOLOGICA BELGICA                                                01075001        INSERTED
                                                                              01075002        INSERTED
SCATTERED HOLDINGS 1961-64                                                FR* 01075003        INSERTED
                                                                      WO      01075004        INSERTED
ACTA ANAESTHESIOLOGICA SCANDINAVICA                                           01100001        INSERTED
                                                                         +00601100002         INSERTED
1-      1957-                                                               * 01100003        INSERTED
SUPPLEMENTS 1-      1959-                                             WO      01100004        INSERTED
ACTA ANATOMICA                                                                01125001        INSERTED
                                                                         +04401125002         INSERTED
1-      1945-                                                               * 01125003        INSERTED
SUPPLEMENTS 1-      1944-                                             QS      01125004        INSERTED
                                                                              99999991        INSERTED
                                                                              99999992        INSERTED
                                                                              99999993        INSERTED
                                                                              99999994        INSERTED

$ENDRUN

    500104SILENCE   10                          UPDATE SUMMARY                                    12/18/70   PAGE  5

SUMMARY OF RECORDS PROCESSED

INPUT FILE. RECORDS READ                        0
  BCD RECORDS                 0
  BIN RECORDS                 0

TRANSACTION FILE ACTIVITY
 INSERTIONS                                   184
  BCD RECORDS              184
  BIN RECORDS                0

 DELETIONS                                      0
  BCD RECORDS                 0
  BIN RECORDS                 0

TOTAL RECORDS ON OUTPUT FILE
INPUT + INSERTIONS - DELETIONS                184  COMPUTED

TOTAL RECORDS ON OUTPUT FILE                  184  COUNTED
  BCD RECORDS              184
  BIN RECORDS                0
```

000 0140
(b)

000 0140
(b)

SAMPLE OUTPUT——HOLD

```
                                                                NOVEMBER 24,1970
ABT                                                                  2500    1
                                                    X            -0  2500    2
      ABSTRACTS OF BIOANALYTIC TECHNOLOGY                   PO       2500    3
                                                                     2500    4


ACM                                                                  5000    1
                                                    X            -0  5000    2
      ASSOCIATION FOR COMPUTING MACHINERY                            5000    3
                                                                     5000    4


AIBS BULLETIN                                                        7500    1
                                                    X            -0  7500    2
      BIOSCIENCE                                                     7500    3
                                                                     7500    4


AMA AMERICAN JOURNAL OF DISEASES OF CHILDREN                        10000    1
                                                    X            -0 10000    2
      AMERICAN JOURNAL OF DISEASES OF CHILDREN                      10000    3
                                                                    10000    4


AMA ARCHIVES ...                                                    12500    1
                                                    X            -0 12500    2
      ARCHIVES ...                                                  12500    3
                                                                    12500    4


AMA JOURNAL OF DISEASES OF CHILDREN                                 15000    1
                                                    X            -0 15000    2
      AMERICAN JOURNAL OF DISEASES OF CHILDREN                      15000    3
                                                                    15000    4


AMA NEWS                                                            17500    1
                                                            +    -0 17500    2
      LIBRARY RETAINS CURRENT YEAR ONLY                             17500    3
                                                    W               17500    4


AORN JOURNAL                                                        20000    1
                                                            +     3 20000    2
      1-      1963-                                     N      *    20000    3
                                                    WY              20000    4

ASHA MONOGRAPHS                                                     22500    1
                                                            +     7 22500    2
      1-      1950-                                            *    22500    3
                                                    WV              22500    4


ASHA REPORTS                                                        25000    1
                                                            +     1 25000    2
      1-      1965-                                                 25000    3
                                                    WV              25000    4
```

continued

ABHANDLUNGEN AUS DEM GESAMTGEBIETE DER MEDIZIN					27500	1
				1	27500	2
1924			GE		27500	3
	WB				27500	4
ABHANDLUNGEN AUS DER AUGENHEILKUNDE UND IHREN GRENZGEBIETEN					30000	1
	X			-0	30000	2
BIBLIOTHECA OPHTHALMOLOGICA					30000	3
					30000	4
ABHANDLUNGEN AUS DER KINDERHEILKUNDE UND IHREN GRENZGEBIETEN					32500	1
	X			-0	32500	2
BIBLIOTHECA PAEDIATRICA					32500	3
					32500	4
ABHANDLUNGEN AUS DER NEUROLOGIE, PSYCHIATRIE, PSYCHOLOGIE UND					35000	1
IHREN GRENZGEBIETEN				1	35000	2
42 1927		G	E		35000	3
	WLWM				35000	4
ABSTRACTS OF BACTERIOLOGY					37500	1
				9	37500	2
1-9 1917-25	A				37500	3
	QW				37500	4
ABSTRACTS OF BIOANALYTIC TECHNOLOGY					40000	1
		S	+	5	40000	2
1-13 1953-65	A				40000	3
	QY				40000	4
ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE					42500	1
	X			-0	42500	2
U.S. PUBLIC HEALTH SERVICE. DIV. OF VENEREAL DISEASE.					42500	3
ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE					42500	4
ABSTRACTS OF HOSPITAL MANAGEMENT STUDIES					45000	1
			+	1	45000	2
1- 1965-	A				45000	3
	WX				45000	4
ABSTRACTS OF HUMAN DEVELOPMENTAL BIOLOGY					47500	1
	X			-0	47500	2
EXCERPTA MEDICA. SECTION 21. DEVELOPMENTAL BIOLOGY AND					47500	3
TERATOLOGY					47500	4
ABSTRACTS OF JAPANESE MEDICINE					50000	1
				6	50000	2
1-2 1960-62	A				50000	3
	WB				50000	4

continued

000 0140
(c)

```
ABSTRACTS OF MYCOLOGY                                                            52500    1
                                                                  +     -0       52500    2
      (1)-      1967-                                      A                     52500    3
                                                           QW                    52500    4


ABSTRACTS OF SOVIET MEDICINE                                                     55000    1
                                                                         2       55000    2
      5       1961                                         A                     55000    3
                                                           WB                    55000    4


ABSTRACTS OF SOVIET MEDICINE. CANCER RESEARCH                                    57500    1
                                                                         1       57500    2
      5-9       1953-56                                    A                     57500    3
                                                           QZ                    57500    4


ABSTRACTS OF SOVIET MEDICINE. PART A. BASIC MEDICAL SCIENCES                     60000    1
                                                                         4       60000    2
      1-4       1957-60                                    A                     60000    3
                                                           W                     60000    4


ABSTRACTS OF SOVIET MEDICINE. PART B. CLINICAL MEDICINE                          62500    1
                                                                         6       62500    2
      1-4       1957-60                                    A                     62500    3
                                                           WB                    62500    4
                .                        .                                         .
                .                        .                                         .
                .                        .                                         .
ACTA ALLERGOLOGICA                                                              105000    1
                                                                  +     13      105000    2
      1-        1948-                                                 *         105000    3
      SUPPLEMENTS 1-7       1950-60                        W3                   105000    4


ACTA ANAESTHESIOLOGICA BELGICA                                                  107500    1
                                                                        -0      107500    2
      SCATTERED HOLDINGS 1961-64                                  FR  *         107500    3
                                                           WO                   107500    4


ACTA ANAESTHESIOLOGICA SCANDINAVICA                                             110000    1
                                                                  +      6      110000    2
      1-        1957-                                                 *         110000    3
      SUPPLEMENTS 1-       1959-                           WO                   110000    4


ACTA ANATOMICA                                                                  112500    1
                                                                  +     44      112500    2
      1-        1945-                                                 *         112500    3
      SUPPLEMENTS 1-       1944-                           OS                   112500    4
```

continued

000 0140
(c)

NOVEMBER 24,1970

	NUMBER OF TITLES	NUMBER OF VOLUMES
NURSING	1	3
CURRENTLY RECEIVED	16	299
INDEXES, ABSTRACTS OR BIBLIOGRAPHIES	11	62
INDEXED IN INDEX MEDICUS	10	269
INDEXED IN I.N.I.	0	0
MICROFORM	0	0
DANISH	0	0
DUTCH	0	0
FRENCH	7	368
GERMAN	1	1
HEBREW	0	0
ITALIAN	1	5
JAPANESE	0	0
RUSSIAN	0	0
SPANISH	0	0
PORTUGUESE	3	2
SERBO-COATIAN	0	0
SLOVENIAN	0	0
ALBANIAN	0	0
NORWEGIAN	0	0
POLISH	0	0
TOTAL	31	523

SAMPLE OUTPUT—PUBLSH

```
 $COPIES    2   U04

        INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY                1

             COMPLETE LIST, INCLUDING HOLDINGS
                                                   NOVEMBER 24,1970


ABT
      SEE
      ABSTRACTS OF BIOANALYTIC TECHNOLOGY

ACM
      SEE
      ASSOCIATION FOR COMPUTING MACHINERY

AIBS BULLETIN
      SEE
      BIOSCIENCE

AMA AMERICAN JOURNAL OF DISEASES OF CHILDREN
      SEE
      AMERICAN JOURNAL OF DISEASES OF CHILDREN

AMA ARCHIVES ...
      SEE
      ARCHIVES ...

AMA JOURNAL OF DISEASES OF CHILDREN
      SEE
      AMERICAN JOURNAL OF DISEASES OF CHILDREN

+ AMA NEWS
      LIBRARY RETAINS CURRENT YEAR ONLY

+ AORN JOURNAL
N     1-       1963-

+ ASHA MONOGRAPHS
      1-       1950-

+ ASHA REPORTS
      1-       1965-

ABHANDLUNGEN AUS DEM GESAMTGEBIETE DER MEDIZIN
      1924

ABHANDLUNGEN AUS DER AUGENHEILKUNDE UND IHREN GRENZGEBIETEN
      SEE
      BIBLIOTHECA OPHTHALMOLOGICA

ABHANDLUNGEN AUS DER KINDERHEILKUNDE UND IHREN GRENZGEBIETEN
      SEE
      BIBLIOTHECA PAEDIATRICA
ABHANDLUNGEN AUS DER NEUROLOGIE, PSYCHIATRIE, PSYCHOLOGIE UND
      IHREN GRENZGEBIETEN
G     42       1927

ABSTRACTS OF BACTERIOLOGY
      1-9      1917-25

+ ABSTRACTS OF BIOANALYTIC TECHNOLOGY
      INCLUDES SUPPLEMENTS
      1-13      1953-65

ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE
      SEE
      U.S. PUBLIC HEALTH SERVICE. DIV. OF VENEREAL DISEASE.
      ABSTRACTS OF CURRENT LITERATURE ON VENEREAL DISEASE

+ ABSTRACTS OF HOSPITAL MANAGEMENT STUDIES
      1-       1965-
```

continued

ABSTRACTS OF HUMAN DEVELOPMENTAL BIOLOGY
SEE
EXCERPTA MEDICA. SECTION 21. DEVELOPMENTAL BIOLOGY AND
TERATOLOGY

ABSTRACTS OF JAPANESE MEDICINE
1-2 1960-62

+ ABSTRACTS OF MYCOLOGY
(1)- 1967-

ABSTRACTS OF SOVIET MEDICINE
5 1961

ABSTRACTS OF SOVIET MEDICINE. CANCER RESEARCH
5-9 1953-56

ABSTRACTS OF SOVIET MEDICINE. PART A. BASIC MEDICAL SCIENCES
1-4 1957-60

ABSTRACTS OF SOVIET MEDICINE. PART B. CLINICAL MEDICINE
1-4 1957-60

ABSTRACTS OF VITAMIN LITERATURE
SEE
VITAMIN ABSTRACTS

+ ABSTRACTS OF WORLD MEDICINE
1- 1947-
ABSTRACTS OF WORLD SURGERY, OBSTETRICS AND GYNECOLOGY
1-11 1947-52

+ ACADEMIE DE CHIRURGIE, PARIS. MEMOIRES
90- 1964-

ACADEMIE DE MEDECINE, PARIS. BULLETIN
SEE
ACADEMIE NATIONALE DE MEDECINE, PARIS. BULLETIN

ACADEMIE DES SCIENCES, PARIS. COMPTES RENDUS HEBDOMADAIRES
DES SEANCES
SCATTERED HOLDINGS PRIOR TO 188, 1929. 188-261
1929-65

+ ACADEMIE DES SCIENCES, PARIS. COMPTES RENDUS HEBDOMADAIRES DES
SEANCES. SERIES D - SCIENCES NATURELLES
262- 1966-

+ ACADEMIE NATIONALE DE MEDECINE, PARIS. BULLETIN
SERIES 1, 1-36. SERIES 2, 1-18. SERIES 3, 19-(128)-
1836-

ACADEMIE ROYALE DE CHIRURGIE, PARIS. MEMOIRES
SCATTERED HOLDINGS 1769-87

ACADEMIE ROYALE DE CHIRURGIE, PARIS. MEMOIRES SUR LES SUJETS
PROPOSES POUR LES PRIX
1-5, 1770-96. NEW ED. 1-5, 1819

ACADEMIE ROYALE DE CHIRURGIE, PARIS. RECUEIL DES PIECES QUI
ONT CONCOURU POUR LE PRIX
SEE
ACADEMIE ROYALE DE CHIRURGIE, PARIS. MEMOIRES SUR LES
SUJETS PROPOSES POUR LES PRIX

+ ACADEMY OF MEDICINE OF NEW JERSEY. BULLETIN
7- 1961-

ACADEMY OF PSYCHOSOMATIC MEDICINE. PROCEEDINGS
1 1954

ACADEMY OF SCIENCES OF THE USSR
SEE
AKADEMIIA NAUK SSSR. DOKLADY

ACCADEMIA MEDICA
65-69 1950-54

continued

```
+ ACOUSTICAL SOCIETY OF AMERICA. JOURNAL
        INCLUDES SUPPLEMENTS
        32-        1960-

+ ACTA ALLERGOLOGICA
        1-         1948-
        SUPPLEMENTS 1-7        1950-60

  ACTA ANAESTHESIOLOGICA BELGICA
        SCATTERED HOLDINGS 1961-64

+ ACTA ANAESTHESIOLOGICA SCANDINAVICA
        1-         1957-
        SUPPLEMENTS 1-        1959-

+ ACTA ANATOMICA
        1-         1945-
        SUPPLEMENTS 1-        1944-
```

INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY 1

SERIALS IN THE LIBRARY OF PARTICULAR INTEREST

IN THE FIELD OF NURSING, INCLUDING HOLDINGS

NOVEMBER 24,1970

```
AORN JOURNAL
      1-       1963-
```

INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY 1

NOVEMBER 24,1970

INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY 1

SERIALS IN MICROFORM

NOVEMBER 24,1970

continued

```
        INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY                 1

             COMPLETE LIST, INCLUDING HOLDINGS

                                                   NOVEMBER 24,1970

ABT
      SEE
      ABSTRACTS OF BIOANALYTIC TECHNOLOGY

ACM
      SEE
      ASSOCIATION FOR COMPUTING MACHINERY

AIBS BULLETIN
      SEE
      BIOSCIENCE

AMA AMERICAN JOURNAL OF DISEASES OF CHILDREN
      SEE
      AMERICAN JOURNAL OF DISEASES OF CHILDREN
         .                              .
         .                              .
         .                              .
+ ACOUSTICAL SOCIETY OF AMERICA. JOURNAL
       INCLUDES SUPPLEMENTS
       32-      1960-

+ ACTA ALLERGOLOGICA
       1-       1948-
       SUPPLEMENTS 1-7      1950-60

  ACTA ANAESTHESIOLOGICA BELGICA
       SCATTERED HOLDINGS 1961-64

+ ACTA ANAESTHESIOLOGICA SCANDINAVICA
       1-       1957-
       SUPPLEMENTS 1-      1959-

+ ACTA ANATOMICA
       1-       1945-
       SUPPLEMENTS 1-      1944-


        INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY                 1

        SERIALS IN THE LIBRARY OF PARTICULAR INTEREST

         IN THE FIELD OF NURSING, INCLUDING HOLDINGS

                                                   NOVEMBER 24,1970

 AORN JOURNAL
       1-       1963-


        INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY                 1
                                                   NOVEMBER 24,1970


        INDIANA UNIVERSITY SCHOOL OF MEDICINE LIBRARY                 1

        SERIALS IN MICROFORM

                                                   NOVEMBER 24,1970
```

SAMPLE OUTPUT——LANSUB

```
      $COPIES    2   U04

                                                                   1

          PRINT FRENCH SURGERY SERIALS (TEST OUTPUT)


+ ACADEMIE DE CHIRURGIE, PARIS. MEMOIRES

  ACADEMIE ROYALE DE CHIRURGIE, PARIS. MEMOIRES

  ACADEMIE ROYALE DE CHIRURGIE, PARIS. MEMOIRES SUR LES SUJETS
      PROPOSES POUR LES PRIX

  ACTA ANAESTHESIOLOGICA BELGICA


                                                                   1

          PRINT GERMAN - NERVOUS SYSTEM SERIALS (TEST OUTPUT)



                                                                   1

          PRINT FRENCH BIOLOGY AND PHYSIOLOGY (TEST OUTPUT)


+ ACADEMIE DES SCIENCES, PARIS. COMPTES RENDUS HEBDOMADAIRES DES
      SEANCES. SERIES D - SCIENCES NATURELLES



                                                                   1

          PRINT FRENCH SURGERY SERIALS (TEST OUTPUT)


+ ACADEMIE DE CHIRURGIE, PARIS. MEMOIRES

  ACADEMIE ROYALE DE CHIRURGIE, PARIS. MEMOIRES

  ACADEMIE ROYALE DE CHIRURGIE, PARIS. MEMOIRES SUR LES SUJETS
      PROPOSES POUR LES PRIX

  ACTA ANAESTHESIOLOGICA BELGICA


                                                                   1

          PRINT GERMAN - NERVOUS SYSTEM SERIALS (TEST OUTPUT)


                                                                   1

          PRINT FRENCH BIOLOGY AND PHYSIOLOGY (TEST OUTPUT)


+ ACADEMIE DES SCIENCES, PARIS. COMPTES RENDUS HEBDOMADAIRES DES
      SEANCES. SERIES D - SCIENCES NATURELLES
```

COST ESTIMATE

Computer time may be estimated from the information presented below and multiplied by the rate for the IUPUI's IBM 7040 ($90./hr.).

Identification	*Time*
Overhead time	4 min.
UPDATE 1 and sort processing	48 serials/sec.
UPDATE 2 and HOLD	1 serial/sec.
PUBLSH	2 serials/sec.
LANSUB	5 serials/sec. (searching for one subject)

Charge to user = computer costs + network overhead

CONTENTS—UPDATE, HOLD, PUBLSH, LANSUB

pages

1– 3 Identification & Abstract
4– 6 User Instructions—UPDATE 1
7 User Instructions—UPDATE 2
9 User Instructions—HOLD
11–12 User Instructions—PUBLSH
13 User Instructions—LANSUB
15–16 Sample Input
17–18 Sample Output—UPDATE 2
19–22 Sample Output—HOLD
23–26 Sample Output—PUBLSH
27 Sample Output—LANSUB
29 Cost—Contents

DESCRIPTIVE TITLE	Correlation and Orthogonal Factor Analysis
CALLING NAME	UMST550
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Lawrence Liddiard University Computer Center University of Minnesota
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 316 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 376-7232

FUNCTIONAL ABSTRACT

This program performs correlation analysis followed by orthogonal factor analysis. The input is read into the computer using variable format; and the input data may have one of three forms: raw data by observation, lower triangle or a correlation (or pseudo-correlation) matrix, or orthogonal factor matrix. If the input is raw data, there are twenty-one possible methods of transformation and/or generation of individual raw variables. After the above methods, or separately, one of six complete transformations may be made on all of the variables (raw and generated).

The following is always given as printed output.

If raw data input:

> The standard error of a factor loading computed from the corrected formula given in Table B on page 441 of Ref. 1.
>
> A Name Card echo print if Name Cards are used.
>
> A Transgeneration Card echo print if transgeneration is used.
>
> Total number of observations.

continued

If raw data or correlation input:

The type of factor analysis component with the diagonal unchanged, common factor with squared multiple correlation on the diagonal, common factor with highest off diagonal absolute value on the diagonal, image covariance, image correlation or independent scale.

The average absolute value of the off-diagonal elements.

The trace (sum of elements on the diagonal) of the matrix to be factor analyzed.

The standard error of a factor loading times the square root of the number of observations.

The communality estimates (the diagonal of the matrix to be factor analyzed).

The eigenvalues for the principal-axis solution, the eigenvalue percent of the number of variables, accumulative percentage of the number of variables, the eigenvalue percent of the common variance (trace of the matrix), and the accumulative percentage of common variance for the computed number of factors.

For all input:

The Problem Card echo print
The Format Card(s) echo print
A message with respect to the total amount of computer core not used or needed above that alloted.

Any or all of the following may be calculated and given as output:

1. Means, standard deviations, variances, covariances, correlations and raw cross products (not usually chosen as output because of massiveness).
2. Means, standard deviations and the elements of the lower triangle of the correlation matrix.
3. An echo of the input correlation or pseudo-correlation matrix.
4. The means and standard deviations on a separate output file for factor scoring or punch output.
5. The elements of the lower triangle of the image analysis matrix (correlation, covariance or independent scale).
6. The data-to-data transformation matrix of image analysis on a separate output file for factor scoring or punch output.
7. The principal-axis factor-loading matrix for the computed number of factors; bordered on the left by the names

continued

(if supplied), indices and communalities (sum of squared row elements) of the variables; bordered on the bottom by the variances (sum of squared column elements) and percentages of trace (variance*100/trace) of the factors.

8. The sorted principal-axis factor-loading matrix for the computed number of factors; bordered on the bottom by the variances and percentages of trace of the factors. Each column of the factor loading matrix is sorted in descending order with the corresponding variable indices or names given in a column to the left of that factor column.
9. The principal-axis factor-scoring matrix for the computed number of factors; bordered on the left by the names or indices and a line count within a variable set. This same information is also put on a separate output file for factor scoring or punch output.
10. The quartimax factor-loading matrices for a stepped range of factors and bordered as in 7.
11. The sorted quartimax factor-loading matrices for a stepped range of factors and bordered as in 8.
12. The quartimax factor-scoring matrices for a stepped range of factors and bordered as in 9. This same information is also put on a separate output file for factor scoring or punch output.
13. The varimax factor-loading matrices for a stepped range of factors and bordered as in 7.
14. The sorted varimax factor-loading matrices for a stepped range of factors and bordered as in 8.
15. The varimax factor-scoring matrices for a stepped range of factors and bordered as in 9. This same information is placed on a separate output file for factor scoring or punch output.

While UMST550 does not compute factor scores itself, auxillary programs are available which can compute and print those scores for each observation when raw data are input. A user wishing to compute factor scores must store his raw data and relevant intermediate results on scratch tape (TAPE99). When filling out the Problem Card, the user must select those options which put that data onto TAPE99. If a user desires factor score output, a note explaining the methodology a user desires to follow in computing those factor scores must accompany the user's EIN submission.

Limits inherent to UMST550

Total number of variables, $2 \leq NV \leq 250$ (approximately) for a CDC 6600 with a 65K—word core.
Total number of observations if given on Problem Card $2 \leq NOBS \leq 99999$

continued

Number of Variable Format Cards, $1 \leq NFC \leq 9$
Number of Transgeneration Cards $0 \leq NTGC \leq 999$
Number of factors $1 \leq NF \leq NV/3$ unless a specific stepped range is chosen or forced in Cols. 59 and 60 of the Problem Card.

Limits which may be supplied by the user

Lower limit for factor eigenvalue (the principal factor variance)—a value of 1.0 may be supplied by the program
Orthogonal rotation cycle limit—a maximum of 40 or NV/2 cycle limit may be supplied by the program
Orthogonal rotation radian angle limit—.005 radians may be supplied by the program

REFERENCES

1. Harman, H.H., *Modern Factor Analysis* (Chicago: Univ. of Chicago Press, 1960).

2. Anderson, D. and Frisch, M., *Statistical Programs for Use on the CDC 6600 Computer* (Minneapolis, Minn.: Univ. of Minn., Univ. Comp. Ctr., 1969).

3. Caffrey, J., "Invrs" Algorithm 66, *Comm. ACM* (July 1961), p. 322.

4. Horst, P., *Factor Analysis of Data Matrices* (New York: Holt, Rinehart and Winston, Inc., 1965).

5. Hotchkiss, R., "SORT2", a standard sorting routine at the Univ. Comp. Ctr. at the Univ. of Minn. based on T.H. Hibbard's "An Empirical Study of Minimal Storage Sorting," *Comm. ACM, 6,* (1968) pp. 206-212.

6. Ortega, J., "The Givens Householder Method for Symmetric Matrices," in A. Ralston and H. Wilf, *Mathematical Methods for Digital Computers, II,* (New York: John Wiley & Sons, Inc., 1967), pp. 94–115.

7. Wilkinson, J.H., "Householder's Method for Symmetric Matrices," Numerische Mathematik, *4* (1962), pp. 354–361.

8. Wilkinson, J.H., "Calculation of the Eigenvalues of a Symmetric Matrix by the Method of Bisection," Numerische Mathematik, *4* (1962), pp. 362–367.

9. Wilkinson, J.H., "Calculation of the Eigenvectors of a Symmetric Matrix by Inverse Iteration," Numerische Mathematik, *4* (1962), pp. 368–376.

USER INSTRUCTIONS

Program Control Cards

Definitions

BLANK: A zero or no character in that column
DIGIT: A non-zero integer in that column

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8		BLANK: no variable Name Cards are present DIGIT: variable Name Cards are present (see section on Name Cards)
9–11	NV or NVI	If the input is a correlation or orthogonal factor matrix, this is the Number of Variables (NV) to be used in the factor analysis. If the input is raw data, this is the Number of Variables (NVI) to be read from the Input Data Cards, and NVI + K = NV where K is the number in Cols. 21–22. NVI does not include the optional counting/card–ordering variable. (NV > 1; NVI > 0)
12–16	NOBS	BLANK: computer counts observations without checking card ordering. (see Input Data Cards) 1: computer counts observations and checks that the first numeric field of the first card of each observation contains a 1. (see Input Data Cards) THE NUMBER OF OBSERVATIONS: If the computer is to read in exactly NOBS observations, $2 \leq \text{NOBS} \leq 99999$. This field applies to raw data input only and otherwise may be ignored.
17	NFC	The Number of Format Cards, $1 \leq \text{NFC} \leq 9$. If the column is BLANK, NFC is set to 1.

continued

Columns	*Parameter*	*Contents*
18–20	NTGC	The Number of Transgeneration Cards, $0 \leq NTGC \leq 999$. If 0 the columns may be left blank. If NTGC < 0, it is set to 0. This field applies to raw data input only and otherwise may be ignored.
21–22	K	The number of new variables added by generation. See section on Transgeneration Cards. This field applies to raw data input only and otherwise may be ignored.
23		BLANK: no complete transformation is to be used DIGIT: the corresponding following complete transformation is to be applied to the NV variables after all transgeneration instructions for an individual observation are completed. 1: $\log_{10}$ transformation 2: $\log_e$ 3: square root 4: square 5: reciprocal 6: exponential of base e 7–9: are not implemented in UMST550 and will cause a terminating error message. This column applies only to raw data input and will be ignored otherwise
24–25		Input Option –DIGIT: correlation matrix input (or pseudo–correlation). In this case Cols. 12–16, 18–23, 57–58 are assumed to be blank and are ignored. BLANK: raw data by observation input +DIGIT: orthogonal factor matrix input (usually a principal factor matrix). In this case Cols. 12–16, 18–23, 26–27, 29–33 are assumed to be blank and are ignored. The maximum of Cols. 43–45 and 52–54 determines the number of factors to be read in.

continued

Columns	*Parameter*	*Contents*
26–27		Intermediate output option –3: the means and standard deviations, if raw data input was selected, are output to the scratch file, TAPE99. All the output for value –2 is also given. –2: the elements of the lower triangle of the image analysis matrix (if Col. 29 $\geq$ 4) are printed and all the output for value –1. –1: if raw data input, the means, standard deviations, variances, covariances, correlations, and raw cross products are printed. If other than raw data input, then no printed output. *Note:* The negative option is not usually selected because of the massive amount of output. BLANK: no intermediate output +1: if raw data input, the means, standard deviations and elements of the lower triangle of the correlation matrix are printed. If correlation input, an echo of the lower triangle of the input matrix is printed. +2: the elements of the lower triangle of the image analysis matrix (if Col. 29 $\geq$ 4) are printed and all the output for value +1 +3: if raw data input was selected, the means and standard deviations are output to the scratch file, TAPE99. All the output for value +2 is also given.
28		Principal Axis output parameter This parameter along with the ones for Quartimax (Col. 39) and Varimax (Col. 48) output determine the output by a binary decomposition of the parameter DIGIT into the binary power values of 1,2,4, and 8 such that the sum of the values is the DIGIT. Thus the DIGIT 5 decomposes into 1+4, meaning that the outputs for value 1 and 4 will be given. BLANK: the program is used only as a correlation program unless Cols. 24–25 contain +DIGIT.

continued

Columns	*Parameter*	*Contents*
28		1: principal–axis factor–loading matrix for the computed number of factors; bordered on the left by the variable names (if supplied), indices, and communalities and bordered on the bottom by variances and percentages of trace. 2: sorted principal–axis factor–loading matrix for the computed number of factors; bordered on the bottom by the variances and percentages of trace and each factor sorted in descending values with corresponding names or indices in a column to the left. 4: principal–axis factor–scoring matrix for the computed number of factors and bordered on the left by the variable names or indices. This matrix is also put on the scratch file, TAPE99. 8: the principal–axis computations are carried out, but no output results from this value.
29		BLANK: common factor analysis with squared multiple correlations put on the diagonal of the correlation or input matrix. If the inverse of the correlation matrix does not exist, then ones or the original diagonal elements are used. 1: component factor analysis—the diagonal of the correlation or input matrix is left unchanged 2: common factor analysis with squared multiple correlations put on the diagonal of the correlation or input matrix. If the inverse of the correlation matrix does not exist, the highest absolute value off–diagonal row element is placed on the diagonal as an estimate of the communality. 3: common factor analysis with the highest absolute value off–diagonal row elements placed on the diagonal

continued

Columns	*Parameter*	*Contents*
29		as communality estimates. This is used when common factor analysis is desired and the inverse of the correlation matrix is known not to exist. 4: image covariance factor analysis. Used in the total approach to the communality problem when those variables with the highest multiple correlation with the other variables are to receive greater weight in the solution (see p. 365, ref. 1) 5: image correlation factor analysis. Used in the total approach (rather than just the diagonal elements) to the communality problem when all variables are to be given equal weight in the solution (see p. 369, ref. 1) 6–9: independent scale factor analysis. Used in the total approach to the communality problem when any particular scaling of the original or image variables is to be ignored in the solution (see p. 370, ref. 1)
30		BLANK: the range of factors is about that number of factors selected by the criterion of Cols. 31–33. DIGIT: the range of factors is taken from Cols. 39–47 and 48–56.
31–33		<0: the factor criterion for eigenvalues is set to 1.0, the Kaiser criterion. If Col. 29 is punched 1, then the resulting solution will be the Kaiser solution. Note that if Col. 29 is not 1 the resulting solution is not equivalent to the Kaiser solution unless the communality estimates are close to 1. ≥0: the factor criterion for eigenvalues is the number with decimal point between Cols. 31 and 32. Thus 092 in these columns selects the

continued

Columns	*Parameter*	*Contents*
31–33		number of factors, NFACT, to be the number of eigenvalues greater than 0.92. Depending on the number of variables this number might range from .50 to 2.00.
34–36	LCYCLE	BLANK or negative: limit on orthogonal cycles is the maximum of 40 and NV/2 DIGIT: limit on orthogonal cycles, 1 ≤ LCYCLE ≤ 999.
37–38	RAL	BLANK or negative: limit on orthogonal rotation angle is .005 radians or 1/4 degree. DIGIT: rotation angle limit is .001 to .099 radians depending on the number 01 to 99 in these columns. If NV is greater than 49, set RAL to at least 09 = .009 radians = 1/2 degree. This will save time in the orthogonal rotations with only a slight decrease in accuracy. (Note that this angle limit determines that the accuracy in the varimax and quartimax solutions can be at most three places.)
39		Quartimax output parameter (see Col. 28 for discussion of this parameter and its binary decomposition) BLANK: no quartimax output or calculations and Cols. 40–47 may be left blank. 1: the quartimax factor–loading matrices for the stepped range specified in the Cols. 40–47. 2: the sorted quartimax factor–loading matrices for the stepped range specified in Cols. 40–47. 4: the quartimax factor–scoring matrices for the stepped range specified in Cols. 40–47. 8: the quartimax calculations are carried out but no output results from this value.

continued

Columns	*Parameter*	*Contents*
40–42	KQS	The starting value of the number of factors for the quartimax stepped range if Col. 30 is DIGIT. Otherwise KQS + NFACT is the starting value. The resulting value is forced to be greater than 0 and less than NV or the computed number of factors.
43–45	KQF	The final value of the number of factors for the quartimax stepped range if Col. 30 is DIGIT. Otherwise KQF+NFACT is the final value. The resulting value is forced to be greater than 0 and less than NV or the computed number of factors.
46–47	KQD	The delta step value for the quartimax stepped range. It is set to one if negative or BLANK. *Examples:* Col. 30 39 40 41 42 43 44 45 46 47 1 2 0 2 0 0 2 7 0 2 This punching of columns would output the sorted quartimax factor loading matrices for matrices with 20, 22, 24, and 26 factors. Col. 30 39 40 41 42 43 44 45 46 47 3 - 1 0 0 0 0 0 5 This punching of columns would output both the quartimax and sorted quartimax factor–loading matrices for NFACT – 10, NFACT – 5, and NFACT factor matrices. Col. 30 39 40 41 42 43 44 45 46 47 7 This punching of columns would output the quartimax factor–loading, sorted factor–loading and factor–scoring matrices for the computed number of factors (NFACT).
48		Varimax output parameter (see Col. 28 for discussion of this parameter and its binary decomposition). BLANK: no varimax output or calculations and Cols. 49–56 may be left blank.

continued

Columns	*Parameter*	*Contents*
48		1: the varimax factor-loading matrices for the stepped range specified in Cols. 49–56. 2: the sorted varimax factor-loading matrices for the stepped range specified in Cols. 49–56. 4: the varimax factor-scoring matrices for the stepped range specified in Cols. 49–56. 8: the varimax calculations are carried out but no output results from this value.
49–51	KVS	The starting value of the number of factors for the varimax stepped range if Col. 30 is DIGIT. Otherwise KVS + NFACT is the starting value. The resulting value is forced to be greater than 0 and less than NV or the computed number of factors.
52–54	KVF	The final value of the number of factors for the varimax stepped range if Col. 30 is DIGIT. Otherwise KVF + NFACT is the final value. The resulting value is forced to be greater than 0 and less than NV or the computed number of factors.
55–56	KVD	The delta step value for the varimax stepped range. Set to 1 if KVD is negative or 0.
57–58	NEXTRA	The additional cells in the raw data input array X used for the intermediate results of transgenerations (see Transgeneration section). Set to 0 if negative.
59–60	NF	The number of possible factors is forced to be other than NV/3. These columns are used to force complete factor sets NF = NV and to save storage when NV is extremely large but the desired number of factors is much less than NV/3.

continued

Columns	*Parameter*	*Contents*
61–80		Name of Problem, the alphanumeric identification of the problem as set by the user. This identification will appear will appear on most of the output headings.

Format Cards

Input by Observation

Variables are considered to be indexed or numbered by subscript according to the order in which they are read in from the Data Cards. The Format Card(s) must provide for reading in the NVI elements of an observation at one time using X–, E–, and F–fields. In addition if either 0 or 1 are punched in Cols. 12–16 of the Problem Card, the first field read by the Format Card shall be an I–field of at least two columns, which will be used in checking card order and/or counting observations.

Correlation matrix input

The Format Card(s) must provide for reading the longest (or last) row of the lower triangle of elements of the symmetric correlation matrix C_{ij} using X–, E–, or F– fields. The first row has the one element C_{11} to be read, the second row has the two elements C_{21} and C_{22} and the NVIth row has the NVI elements $C_{NVI,1}$, $C_{NVI,2}$, ..., $C_{NVI,NVI}$.
Warning: Because only the lower triangle is read, care must be taken to properly punch the cards. For example, if NVI = 30 and 10 variables can be punched per card, the first 10 cards contain the values for the first 10 rows, the next 20 cards contains the values for the second 10 rows, and the next 30 cards contain the values for the last 10 rows.

Orthogonal factor matrix input

The Format Card(s) must provide for reading the NVI rows of the orthogonal factor matrix OF_{ij} using X–, E–, of F– fields. The number of factors in OF_{ij} is determined by the maximum of KQF and KVF (Cols. 43–45 and 52–54 of Problem Card).

Transgeneration Cards

Each Transgeneration Card specifies the transformation or variable generation desired. Transgenerations are performed sequentially; thus, two variables may be added to create a third variable by one transgeneration and the square root of the new

continued

variable may be done by the next transformation. Cols. 18–20 of the Problem Card determine how many of these cards are to be read in. Violation of restrictions relating to the transformations and transgenerations results in the omission of all variables of the observation for which the violation occurred and an appropriate printout indicating the observation omitted. Note the sum of Cols. 9–11 (NVI) and Cols. 21–22 (K) of the Problem Card may not exceed approximately 250 for a 65K CDC 6600 although intermediate calculations may use up to 99 extra X locations if NEXTRA (Cols. 57–58) is used.

Notation

Transformation:	the mutation of an existing variable with the mutation replacing the existing variable
Generation:	the mutation of an existing variable with the mutation replacing another existing variable or producing an additional variable to the present set of variables.
Transgeneration Card:	the instruction to transform or generate specific variables using specified variables and/or constants.
Complete Transformation:	a specific transformation applied to all of the present set of variables.
x:	the variable prior to transgeneration.
x':	the transformed variable.
c:	a constant.
A,B,:	specified variable numbers, or indices.
*:	indicates multiplication.

Code	*Transgeneration*	*Restriction*
00	$x' = x$	
01	$x' = \sqrt{x}$	$x \geq 0$
02	$x' = \sqrt{x} + \sqrt{x + 1}$	$x \geq 0$
03	$x' = \log_{10} x$	$x > 0$
04	$x' = e^{x}$	$x \leq 740$
05	$x' = \sin^{-1} x$	$-1 \leq x \leq 1$
06	$x' = \log_{e} x$	$x > 0$

continued

Code	*Transgeneration*	*Restriction*
07	$x' = \sin x$	$\lvert x \rvert < 2.2 \times 10^{14}$ x in radians
08	$x' = \cos x$	$\lvert x \rvert < 2.2 \times 10^{14}$ x in radians
09	$x' = 1/x$	$x \neq 0$
10	$x' = x + c$	
11	$x' = x * c$	
12	$x' = x^c$	$x > 0$
13	$x' = c^x$	$c > 0$
14	if $x \geq c$, $s' = 1$, else $x' = 0$	
15	if $x_A \geq x_B$, $x' = 1$, else $x' = 0$	
16	$x' = x_A + x_B$	
17	$x' = x_A - x_B$	
18	$x' = x_A * x_B$	
19	$x' = x_A / x_B$	$x_B \neq 0$
20	$x' = {x_A}^{x_B}$	$x_A > 0$
21	if $x = c$, $x' = 1$, else $x' = 0$	

Examples of Transgeneration

Restrictions to keep in mind: NVI, the number of input variables; $1 \leq NVI$. K, the number of additional variables; $-9 \leq K \leq 99$. NV = NVI+K, the number of variables to be used in the analysis; $2 \leq NV \leq 250$. X(I), the individual variable in transgeneration; $1 \leq I \leq NV + NEXTRA$.

Thus the total number of variables in the factor analysis may be more or less than the number of input variables, but this set must be the compact (no gaps) set $x_1, x_2, \ldots, x_{NV}$.

Example 1:

NVI = 6, K = –1, NV = 5, NEXTRA (Cols. 57–58) = 1

continued

Transgeneration

$x_5 = x_6 + x_5$

Thus the variable x_6 is entered into the factor analysis by way of combination with x_5.

Example 2:

NVI = 10, K = 2, NV = 12, NEXTRA (Cols. 57–58) = 2

Transgenerations

x_{13} = SIN (x_6) note that x_{13} and x_{14} are intermediate variables that will not be used in the analysis since NV = 12.

x_{14} = ALOG10(x_7)

$x_{11} = x_{13}/x_{14}$

$x_{12} = x_7 + x_8 + x_9$ + SQRT (x_6) this would take four Transgeneration Cards

Thus two new variables, x_{11} and x_{12} are entered into the factor analysis.

Columns	*Contents*
1– 3	Variable number to be assigned to x'
4– 5	Transgeneration code
6–10	A–variable number, or variable number of x
11–20	B–variable number, or constant (keypunch decimal place if not at extreme right).

Name Cards

If Col. 8 of the Problem Card had a DIGIT punched in it, then at this point there must be [NV/10] Name Cards. Every Name Card has ten 8–character fields, each field specifying a mnemonic to identify its corresponding variable. Thus for NV = 26 there would be three cards; the first ten variable names specified on the first card, the second ten varible names on the second card and the remaining six variable names specified in the first 6 fields of the third card.

Input Data Cards

The Data Cards shall be punched in accordance with the format given on the Format Card(s).

Input by observations

The NVI elements of an observation are read at one time from a set of cards; each observation starting on a new set of cards.

continued

If a 0 or a 1 is punched in Cols. 12–16 of the Problem Card, there must be a non-negative integer in the first field (an I-field) read by the Format Card, for use in determining when the End of Data Card is read (and to check card ordering if 1). When a 1 is punched, it implies that more than one card is required for each observation, and that the cards for each observation are numbered in the order 1, 2, 3 in this specific I-field. Thus, finding a 2 or 3 in this I-field would mean that some cards were out of order or missing.

Correlation matrix input

The j elements of row j are read at one time from a set of cards; each row starting on a new set of cards. NVI is the total number of input rows.

Orthogonal factor matrix input

The maximum of KQF and KVF factor loadings of a row are read at one time from a set of cards, each row starting on a new set of cards. NVI rows are read into the computer.

End of Data Card(s)

This card(s) is required only if Cols. 12–16 of the Problem Card contain 0 or 1. It must have a right adjusted –1 punched where it will be read by the first field (the I-field) of the Format Card. There must be as many blank cards following this card as is necessary to complete a set of observation cards.

Finish Card

In Cols. 1–6 punch the word FINISH. This card appears after the last problem.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards for a problem set.

Program Control Cards for problem 1
  Problem Card for problem 1
  Format Card(s) for problem 1
  Transgeneration Card(s) (if any) for problem 1
  Name Card(s) (if any) for problem 1
Input Data Cards for problem 1
End of Data Card(s) for problem 1 (if required)
Program Control Cards for problem 2

continued

Input Data Cards for problem 2
End of Data Card(s) for problem 2 (if required)
:
:
Program Control Cards for last problem
Input Data Cards for last problem
End of Data Card(s) for last problem (if required)
Finish Card

Additional information on timing, storage and page estimates, error printouts, punch output, and mathematical procedure can be found in Ref. 2.

REFERENCES

1. Horst, P., *Factor Analysis of Data Matrices* (New York: Holt, Rinehart and Winston, Inc., 1965).

2. Anderson, D. and Frisch, M., *Statistical Programs for Use on the CDC 6600 Computer* (Minneapolis, Minn.: Univ. of Minn., Univ. Comp. Ctr., 1969).

SAMPLE INPUT

```
PROBLEM1  9    01         171  -1              7           ONES,KAISER,NAMES
(I4,9F4.3)
 QUES 44 QUES 45 QUES 47 QUES 49 QUES 51 QUES 54 QUES 96 QUES 99 QUES 13
     128 181 421 506 857 746 280 178 246
     764 740 563-387-293-202 261 281 043
    -030-046 014 147-109-135 640 682 661
    -280-351-326-023-109-186 083 091-654
    -336-306-429-542-006-153-428 056-124
    -276-324-271-370-225 035 129-446-124
     057-070-016 006 152-441-166-354-033
    -010-140 326-004-258-091-410-200-101
    -303 227-014 029-102-125-086-106-074
     086 057-003 161 073 134-082-067-009
     164 106-124 234-002 227-072 025 156
     036-074-142 242 023 192-148-119 011
 -1
PROBLEM   9    121         101                 3  2  4 1   SMC, SPECIFIC RANGE
(4X,9F4.3)
     128 181 421 506 857 746 280 178 246
     764 740 563-387-293-202 261 281 043
    -030-046 014 147-109-135 640 682 661
    -280-351-326-023-109-186 083 091-654
    -336-306-429-542-006-153-428 056-124
    -276-324-271-370-225 035 129-446-124
     057-070-016 006 152-441-166-354-033
    -010-140 326-004-258-091-410-200-101
    -303 227-014 029-102-125-086-106-074
     086 057-003 161 073 134-082-067-009
     164 106-124 234-002 227-072 025 156
     036-074-142 242 023 192-148-119 011
PROBLEM   9    11        83        7           3           RAW DATA, MAX OFFDIA
(I4,9F4.3)
   1 128 181 421 506 857 746 280 178 246
   1 764 740 563-387-293-202 261 281 043
   1-030-046 014 147-109-135 640 682 661
   1-280-351-326-023-109-186 083 091-654
   1-336-306-429-542-006-153-428 056-124
   1-276-324-271-370-225 035 129-446-124
   1 057-070-016 006 152-441-166-354-033
   1-010-140 326-004-258-091-410-200-101
   1-303 227-014 029-102-125-086-106-074
   1 086 057-003 161 073 134-082-067-009
   1 164 106-124 234-002 227-072 025 156
   1 036-074-142 242 023 192-148-119 011
 -1
PROBLEM   9     1      -1  74  05                  3          9IMAGE COV,FACT FORCE
(9F4.3)
1000
08291000
076807751000
010801150272100 0
0033006102050636100 0
010801250238062607091000
029803230296024901380190100 0
030903470271018300910103065410 00
0351036903850369025402910527054110 00
PROBLEM   9     1      -1 315  -1                  3 -1  1 1  4IMAGE COR,EIG RANGE
(9F4.3)
1000
08291000
076807751000
010801150272100 0
0033006102050636100 0
010801250238062607091000
0298032302960249013801901000
030903470271018300910103065410 00
0351036903850369025402910527054110 00
PROBLEM   9     1      -1 216  -1       5          7           INDEP SCALE,BINARY,
(9F4.3)
1000
08291000
076807751000
010801150272100 0
0033006102050636100 0
010801250238062607091000
0298032302960249013801901000
030903470271018300910103065410 00
0351036903850369025402910527054110 00
FINISH
```

605

SAMPLE OUTPUT

```
FACTOR ANALYSIS                                      ONES.KAISER.NAMES

TOTAL VARIABLES=   9, INPUT OPTION= -0, INTERMEDIATE OUTPUT=  1, TRANSFORM=-0, NSMC= 1, SELECT RANGE=-0

LIMITS--EIGENVALUE= 1.00, ROTATION CYCLES=  40, ROTATION ANGLE IN RADIANS= .005

PRINCIPAL-FACTOR= 7    QUARTIMAX=-0     -0,  -0,  1     VARIMAX= 7     -0,  -0,  1

EXTRA SPACE IN X ARRAY FOR TRANSGENERATION =     0 MAXIMUM NUMBER OF FACTORS =   3

INPUT DATA FORMAT  (I4,9F4.3)

$$$ SAVE CORE SPACE. UMST550 CAN USE 070200B LESS FIELD LENGTH ON JOB CARD TO RUN THIS PROBLEM
NAME CARD ECHO
 QUES 44 QUES 45 QUES 47 QUES 49 QUES 51 QUES 54 QUES 86 QUES 99 QUES 13


TOTAL OBSERVATIONS=     12




VARIABLE NAME          I         AVERAGE(I)      STANDARD DEVIATION(I)

        QUES 44        1    -4.0708177570E-16       .3013904143
        QUES 45        2    -7.7715611724E-16       .3015535532
        QUES 47        3    -8.3333333334E-05       .3015778079
        QUES 49        4    -8.3333333333E-05       .3013214708
        QUES 51        5     8.3333333333E-05       .3016100608
        QUES 54        6     8.3333333333E-05       .3014376035
        QUES 86        7     8.3333333332E-05       .3014659511
        QUES 99        8     1.7500000000E-03       .2987690275
        QUES 13        9    -1.6666666667E-04       .3013692490



 CORRELATION MATRIX                                  ONES.KAISER.NAMES

     1   1.00000

     2    .82890    1.00000

     3    .76832     .77490    1.00000

     4    .10800     .11510     .27193    1.00000

     5    .03268     .06085     .20493     .63605    1.00000

     6    .10842     .12526     .23857     .62586     .70942    1.00000

     7    .29768     .32257     .29575     .24957     .13827     .18991    1.00000

     8    .30671     .34365     .26819     .17775     .08766     .10525     .66303    1.00000

     9    .35058     .36860     .38489     .36932     .25401     .29139     .52740     .54364    1.00000
COMPONENT FACTOR ANALYSIS - DIAGONAL UNCHANGED

AVERAGE CORRELATION IS     .3373609942     TRACE OF FACTOR ANALYSIS MATRIX IS     9.000000000
S.E. OF A LOADING TIMES SQUARE ROOT OF THE NO. OF OBSERVATIONS IS     1.189642409
STANDARD ERROR OF A FACTOR LOADING IS     .3434201826

COMMUNALITY ESTIMATES(IE DIAGONAL OF MATRIX TO BE FACTOR ANALYZED)
  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000
```

continued

```
COMPUTED NUMBER OF FACTORS =  3  RANK OF CORRELATION INVERSE =   0

                          0/0 OF NUMBER    ACCUMULATIVE     0/0 OF COMMON       ACCUMULATIVE 0/0
   I    EIGENVALUES       OF VARIABLES     PERCENTAGE       VARIANCE(TRACE)     COMMON VARIANCE

   1      3.748546          41.65            41.65             41.65               41.65
   2      2.049715          22.77            64.43             22.77               64.43
   3      1.340619          14.90            79.32             14.90               79.32


        PRINCIPAL FACTOR LOADING MATRIX   WITH   3 FACTORS AND      0 CYCLES    ONES.KAISER.NAMES
  QUES 44   1     .879   -.716   -.492   -.352

  QUES 45   2     .879   -.739   -.478   -.324

  QUES 47   3     .850   -.772   -.295   -.408

  QUES 49   4     .734   -.555    .649   -.070

  QUES 51   5     .801   -.463    .745   -.180

  QUES 54   6     .784   -.519    .693   -.186

  QUES 86   7     .766   -.642   -.081    .589

  QUES 99   8     .798   -.615   -.168    .626

  QUES 13   9     .648   -.716    .033    .367

 VARIANCE  10    7.139   3.749   2.050   1.341

 0/0TRACE  10   79.321  41.651  22.775  14.896



 SORTED PRINCIPAL FACTOR LOADING MATRIX   WITH   3 FACTORS AND      0 CYCLES    ONES.KAISER.NAMES
     QUES 51  -.463     QUES 51   .745     QUES 99   .626

     QUES 54  -.519     QUES 54   .693     QUES 86   .589

     QUES 49  -.555     QUES 49   .649     QUES 13   .367

     QUES 99  -.615     QUES 13   .033     QUES 49  -.070

     QUES 86  -.642     QUES 86  -.081     QUES 51  -.180

     QUES 13  -.716     QUES 99  -.168     QUES 54  -.186

     QUES 44  -.716     QUES 47  -.295     QUES 45  -.324

     QUES 45  -.739     QUES 45  -.478     QUES 44  -.352

     QUES 47  -.772     QUES 44  -.492     QUES 47  -.408

    VARIANCE  3.749    VARIANCE  2.050    VARIANCE  1.341

    0/0TRACE 41.651    0/0TRACE 22.775    0/0TRACE 14.896


 PRINCIPAL FACTOR SCORING MATRIX  3 FACTORS,    9 VARIABLES  ONES.KAISER.NAMES
 QUES 44 1  -.191  -.240  -.263
 QUES 45 1  -.197  -.233  -.242
 QUES 47 1  -.206  -.144  -.304
 QUES 49 1  -.148   .317  -.052
 QUES 51 1  -.123   .363  -.135
 QUES 54 1  -.138   .338  -.139
 QUES 86 1  -.171  -.039   .439
 QUES 99 1  -.164  -.082   .467
 QUES 13 1  -.191   .016   .274
```

continued

```
          VARIMAX FACTOR LOADING MATRIX   WITH   3 FACTORS AND      7 CYCLES     ONES.KAISER.NAMES
  QUES 44   1     .879    -.920     .003    .180
  QUES 45   2     .879    -.913     .020    .214
  QUES 47   3     .850    -.886     .210    .146
  QUES 49   4     .734    -.071     .832    .193
  QUES 51   5     .801    -.020     .894    .040
  QUES 54   6     .784    -.086     .878    .074
  QUES 86   7     .766    -.150     .104    .856
  QUES 99   8     .798    -.157     .009    .879
  QUES 13   9     .648    -.260     .287    .706
 VARIANCE  10    7.139    2.592    2.400   2.147
 0/0TRACE  10   79.321   28.798   26.662  23.861
```

```
SORTED   VARIMAX FACTOR LOADING MATRIX   WITH   3 FACTORS AND      7 CYCLES     ONES.KAISER.NAMES
     QUES 51  -.020     QUES 51   .894    QUES 99   .879
     QUES 49  -.071     QUES 54   .878    QUES 86   .856
     QUES 54  -.086     QUES 49   .832    QUES 13   .706
     QUES 86  -.150     QUES 13   .287    QUES 45   .214
     QUES 99  -.157     QUES 47   .210    QUES 49   .193
     QUES 13  -.260     QUES 86   .104    QUES 44   .180
     QUES 47  -.886     QUES 45   .020    QUES 47   .146
     QUES 45  -.913     QUES 99   .009    QUES 54   .074
     QUES 44  -.920     QUES 44   .003    QUES 51   .040
   VARIANCE  2.592   VARIANCE  2.400   VARIANCE  2.147
   0/0TRACE 28.798   0/0TRACE 26.662   0/0TRACE 23.861
```

```
   VARIMAX FACTOR SCORING MATRIX  3 FACTORS,    9 VARIABLES  ONES.KAISER.NAMES
QUES 44 1  -.393  -.053  -.077
QUES 45 1  -.382  -.049  -.057
QUES 47 1  -.376   .045  -.113
QUES 49 1   .037   .351   .009
QUES 51 1   .033   .398  -.077
QUES 54 1   .008   .385  -.069
QUES 86 1   .103  -.051   .459
QUES 99 1   .101  -.097   .482
QUES 13 1   .029   .043   .330
```

continued

```
FACTOR ANALYSIS                                   SMC, SPECIFIC RANGE

TOTAL VARIABLES=   9, INPUT OPTION= -0, INTERMEDIATE OUTPUT= -0, TRANSFORM=-0, NSMC= 0, SELECT RANGE= 1

LIMITS--EIGENVALUE= 0.00, ROTATION CYCLES=  40, ROTATION ANGLE IN RADIANS= .005

PRINCIPAL-FACTOR= 1    QUARTIMAX=-0    -0,  -0,  1    VARIMAX= 3     2,   4,  1

EXTRA SPACE IN X ARRAY FOR TRANSGENERATION =     0 MAXIMUM NUMBER OF FACTORS =   4

INPUT DATA FORMAT  (4X,9F4.3)

$$$ SAVE CORE SPACE. UMST550 CAN USE 070100B LESS FIELD LENGTH ON JOB CARD TO RUN THIS PROBLEM
UOFM TIME    10.924 SECONDS CORRELATION

TOTAL OBSERVATIONS=      12
UOFM TIME    10.927 SECONDS INVERSE
COMMON FACTOR ANALYSIS - SQUARED MULTIPLE CORRELATIONS PUT ON DIAGONAL

AVERAGE CORRELATION IS      .3373609942     TRACE OF FACTOR ANALYSIS MATRIX IS      5.249243191
S.E. OF A LOADING TIMES SQUARE ROOT OF THE NO. OF OBSERVATIONS IS      1.189642409
STANDARD ERROR OF A FACTOR LOADING IS      .3434201826

COMMUNALITY ESTIMATES(IE DIAGONAL OF MATRIX TO BE FACTOR ANALYZED)
   .73324    .74374    .69073    .51119    .57249    .56209    .49006    .50833    .43738



COMPUTED NUMBER OF FACTORS =  4  RANK OF CORRELATION INVERSE =   9

                      0/0 OF NUMBER    ACCUMULATIVE    0/0 OF COMMON      ACCUMULATIVE 0/0
  I   EIGENVALUES     OF VARIABLES     PERCENTAGE      VARIANCE(TRACE)    COMMON VARIANCE

  1     3.353006          37.26           37.26             63.88              63.88
  2     1.650511          18.34           55.59             31.44              95.32
  3      .895780           9.95           65.55             17.06             112.38
  4     -.031926           -.35           65.19              -.61             111.78



        PRINCIPAL FACTOR LOADING MATRIX   WITH   4 FACTORS AND     0 CYCLES    SMC, SPECIFIC RANGE

             1     .791    -.721    -.453    -.255    -.024

             2     .805    -.745    -.439    -.230    -.064

             3     .749    -.763    -.256    -.307     .083

             4     .588    -.497     .576    -.066     .076

             5     .659    -.419     .671    -.175    -.044

             6     .646    -.472     .625    -.170    -.067

             7     .579    -.579    -.016     .493    -.029

             8     .605    -.558    -.090     .534    -.035

             9     .509    -.637     .073     .303     .078

VARIANCE    10    5.931    3.353    1.651     .896     .032

0/0TRACE    10  112.992   63.876   31.443   17.065     .608
```

continued

```
         VARIMAX FACTOR LOADING MATRIX   WITH   2 FACTORS AND     7 CYCLES    SMC, SPECIFIC RANGE
              1     .725    -.850   -.056
              2     .748    -.864   -.032
              3     .647    -.793    .137
              4     .578    -.163    .743
              5     .627    -.050    .790
              6     .613    -.118    .774
              7     .335    -.517    .262
              8     .319    -.533    .186
              9     .411    -.525    .367
VARIANCE     10    5.004    2.968   2.036
0/0TRACE     10   95.319   56.534  38.784

                    .                  .
                    .                  .
                    .                  .

         VARIMAX FACTOR LOADING MATRIX   WITH   4 FACTORS AND     1 CYCLES    SMC, SPECIFIC RANGE
              1     .791    -.865    .009    .204   -.017
              2     .805    -.863    .027    .238   -.056
              3     .749    -.816    .207    .175    .094
              4     .588    -.077    .731    .196    .096
              5     .659    -.027    .809    .060   -.023
              6     .646    -.086    .792    .094   -.046
              7     .579    -.172    .122    .731   -.019
              8     .605    -.176    .039    .756   -.027
              9     .509    -.258    .271    .601    .091
VARIANCE     10    5.931    2.300   1.951   1.647    .034
0/0TRACE     10  112.992   43.823  37.159  31.368    .642


SORTED   VARIMAX FACTOR LOADING MATRIX   WITH   4 FACTORS AND     1 CYCLES    SMC, SPECIFIC RANGE
      5  -.027      5   .809      8   .756      4   .096
      4  -.077      6   .792      7   .731      3   .094
      6  -.086      4   .731      9   .601      9   .091
      7  -.172      9   .271      2   .238      1  -.017
      8  -.176      3   .207      1   .204      7  -.019
      9  -.259      7   .122      4   .196      5  -.023
      3  -.816      8   .039      3   .175      8  -.027
      2  -.863      2   .027      6   .094      6  -.046
      1  -.865      1   .009      5   .060      2  -.056
     10  2.300     10  1.951     10  1.647     10   .034
     10 43.823     10 37.159     10 31.368     10   .642
```

continued

610

```
FACTOR ANALYSIS                                        RAW DATA, MAX OFFDIA

TOTAL VARIABLES=   9, INPUT OPTION= -0, INTERMEDIATE OUTPUT= -0, TRANSFORM=-0, NSMC= 3, SELECT RANGE=-0

LIMITS--EIGENVALUE=-0.00, ROTATION CYCLES=  40, ROTATION ANGLE IN RADIANS= .005

PRINCIPAL-FACTOR= 8    QUARTIMAX= 7    -0,  -0,  1    VARIMAX= 3    -0,  -0,  1

EXTRA SPACE IN X ARRAY FOR TRANSGENERATION =     0 MAXIMUM NUMBER OF FACTORS =   3

INPUT DATA FORMAT  (I4,9F4.3)

$$$ SAVE CORE SPACE. UMST550 CAN USE 070100B LESS FIELD LENGTH ON JOB CARD TO RUN THIS PROBLEM
UOFM TIME    11.241 SECONDS CORRELATION

TOTAL OBSERVATIONS=     12
COMMON FACTOR ANALYSIS - HIGHEST MAGNITUDE OFF-DIAGONAL IN ROW PUT ON DIAGONAL

AVERAGE CORRELATION IS     .3373609942     TRACE OF FACTOR ANALYSIS MATRIX IS     6.357270537
S.E. OF A LOADING TIMES SQUARE ROOT OF THE NO. OF OBSERVATIONS IS     1.189642409
STANDARD ERROR OF A FACTOR LOADING IS     .3434201826

COMMUNALITY ESTIMATES(IE DIAGONAL OF MATRIX TO BE FACTOR ANALYZED)
   .82890    .82890    .77490    .63605    .70942    .70942    .66303    .66303    .54364



COMPUTED NUMBER OF FACTORS =  3  RANK OF CORRELATION INVERSE =    0

                        0/0 OF NUMBER     ACCUMULATIVE     0/0 OF COMMON       ACCUMULATIVE 0/0
  I    EIGENVALUES      OF VARIABLES      PERCENTAGE       VARIANCE(TRACE)     COMMON VARIANCE

  1     3.467578          38.53             38.53             54.55                54.55
  2     1.774331          19.71             58.24             27.91                82.46
  3     1.033636          11.48             69.73             16.26                98.71



          QUARTIMAX FACTOR LOADING MATRIX   WITH   3 FACTORS AND     7 CYCLES    RAW DATA, MAX OFFDIA

               1     .826    -.902    -.005    .111

               2     .828    -.898     .014    .146

               3     .767    -.849     .198    .083

               4     .614    -.107     .759    .164

               5     .707    -.043     .839    .025

               6     .695    -.107     .825    .053

               7     .651    -.236     .135    .760

               8     .670    -.243     .049    .780

               9     .517    -.322     .288    .575

VARIANCE      10    6.276    2.584    2.104    1.588

0/0TRACE      10   98.714   40.644   33.098   24.972


                     .                          .
                     .                          .
                     .                          .
```

continued

```
SORTED   VARIMAX FACTOR LOADING MATRIX   WITH   3 FACTORS AND     7 CYCLES   RAW DATA, MAX OFFDIA
     5  -.024     5   .839     8   .800
     4  -.077     6   .825     7   .781
     6  -.085     4   .756     9   .609
     7  -.165     9   .278     2   .227
     8  -.171     3   .211     4   .193
     9  -.263     7   .119     1   .192
     3  -.834     8   .032     3   .164
     2  -.881     2   .025     6   .083
     1  -.888     1   .008     5   .050
    10  2.400    10  2.093    10  1.782
    10 37.754    10 32.926    10 28.035
```

```
FACTOR ANALYSIS                                     IMAGE COV,FACT FORCE
TOTAL VARIABLES=   9, INPUT OPTION= -1, INTERMEDIATE OUTPUT= -0, TRANSFORM=-0, NSMC= 4, SELECT RANGE=-0

LIMITS--EIGENVALUE=  .05, ROTATION CYCLES=  40, ROTATION ANGLE IN RADIANS= .005

PRINCIPAL-FACTOR= 7    QUARTIMAX=-0    -0,  -0,  1    VARIMAX= 3    -0,  -0,  1

EXTRA SPACE IN X ARRAY FOR TRANSGENERATION =    0 MAXIMUM NUMBER OF FACTORS =   9

INPUT DATA FORMAT  (9F4.3)

$$$ SAVE CORE SPACE. UMST550 CAN USE 070000B LESS FIELD LENGTH ON JOB CARD TO RUN THIS PROBLEM
UOFM TIME    11.484 SECONDS INVERSE
IMAGE COVARIANCE FACTOR ANALYSIS

AVERAGE CORRELATION IS      .2922614946     TRACE OF FACTOR ANALYSIS MATRIX IS      5.229363373
S.E. OF A LOADING TIMES SQUARE ROOT OF THE NO. OF OBSERVATIONS IS      1.336519719

COMMUNALITY ESTIMATES(IE DIAGONAL OF MATRIX TO BE FACTOR ANALYZED)
   .73307    .74435    .69053    .51136    .57201    .56242    .48081    .49894    .43586


COMPUTED NUMBER OF FACTORS =  6  RANK OF CORRELATION INVERSE =   9

                       0/0 OF NUMBER    ACCUMULATIVE     0/0 OF COMMON      ACCUMULATIVE 0/0
  I   EIGENVALUES     OF VARIABLES     PERCENTAGE       VARIANCE(TRACE)    COMMON VARIANCE

  1      3.013706         33.49            33.49              57.63              57.63
  2      1.333692         14.82            48.30              25.50              83.13
  3       .579224          6.44            54.74              11.09              94.21
  4       .125416          1.39            56.13               2.40              96.61
  5       .065089           .72            56.86               1.24              97.85
  6       .050784           .56            57.42                .97              98.82
```

continued

```
PRINCIPAL FACTOR LOADING MATRIX   WITH   6 FACTORS AND     0 CYCLES    IMAGE COV.FACT FORCE
          1     .718   -.713   -.399   -.189    .029   -.116    .009

          2     .740   -.733   -.383   -.171    .050    .123   -.096

          3     .682   -.748   -.233   -.235   -.076   -.014    .081

          4     .505   -.449    .531   -.059   -.012    .046   -.125

          5     .566   -.377    .604   -.145    .146   -.129   -.016

          6     .560   -.429    .562   -.147   -.155    .081    .094

          7     .475   -.525    .015    .404    .165    .063    .063

          8     .498   -.508   -.054    .430   -.202   -.084   -.066

          9     .424   -.589    .088    .254    .054    .009    .045

VARIANCE 10    5.168   3.014   1.334    .579    .125    .065    .051

0/0TRACE 10   98.825  57.630  25.504  11.076   2.398   1.245    .971

                 .                       .                       .
                 .                       .                       .
                 .                       .                       .

SORTED   VARIMAX FACTOR LOADING MATRIX   WITH   6 FACTORS AND    17 CYCLES    IMAGE COV.FACT FORCE
     5  -.029     5    .722     7    .631     7   .145     6    .202     5    .089

     4  -.079     6    .707     8    .615     5   .073     3    .091     1    .078

     6  -.086     4    .672     9    .530     2   .037     8    .027     3    .068

     7  -.194     9    .263     2    .237     9   .035     9    .004     6    .031

     8  -.205     3    .192     1    .212     1  -.033     2    .001     9    .017

     9  -.269     7    .131     3    .192     4  -.039     7   -.021     8   -.003

     3  -.770     8    .065     4    .179     3  -.044     4   -.024     7   -.016

     2  -.809     2    .044     6    .099     6  -.047     1   -.073     4   -.110

     1  -.813     1    .023     5    .081     8  -.270     5   -.158     2   -.163

    10  2.074    10   1.602    10   1.244    10   .108    10    .081    10    .059

    10 39.666    10  30.632    10  23.784    10  2.070    10   1.550    10   1.123
```

continued

10

```
FACTOR ANALYSIS                                         IMAGE COR.EIG RANGE
TOTAL VARIABLES=   9, INPUT OPTION= -1, INTERMEDIATE OUTPUT=  3, TRANSFORM=-0, NSMC= 5, SELECT RANGE=-0
LIMITS--EIGENVALUE= 1.00, ROTATION CYCLES=  40, ROTATION ANGLE IN RADIANS= .005
PRINCIPAL-FACTOR= 1    QUARTIMAX=-0    -0,  -0,  1    VARIMAX= 3    -1,   1,  1
EXTRA SPACE IN X ARRAY FOR TRANSGENERATION =    0 MAXIMUM NUMBER OF FACTORS =   4
INPUT DATA FORMAT  (9F4.3)
$$$ SAVE CORE SPACE. UMST550 CAN USE 070100B LESS FIELD LENGTH ON JOB CARD TO RUN THIS PROBLEM


 CORRELATION MATRIX                                     IMAGE COR.EIG RANGE

     1   1.00000
     2    .82900   1.00000
     3    .76800    .77500   1.00000
     4    .10800    .11500    .27200   1.00000
     5    .03300    .06100    .20500    .63600   1.00000
     6    .10800    .12500    .23800    .62600    .70900   1.00000
     7    .29800    .32300    .29600    .24900    .13800    .19000   1.00000
     8    .30900    .34700    .27100    .18300    .09100    .10300    .65400   1.00000
     9    .35100    .36900    .38500    .36900    .25400    .29100    .52700    .54100   1.00000
IMAGE CORRELATION FACTOR ANALYSIS


 IMAGE OR INDEPENDENT SCALE MATRIX                               IMAGE COR.EIG RANGE

     1   1.00000
     2    .93141   1.00000
     3    .93057    .92888   1.00000
     4    .19663    .23867    .36119   1.00000
     5    .10076    .10292    .26946    .90147   1.00000
     6    .15740    .18033    .38338    .92143    .85655   1.00000
     7    .49461    .52869    .49924    .43998    .30721    .30815   1.00000
     8    .49895    .50330    .51918    .35324    .15240    .26744    .80933   1.00000
     9    .59362    .62791    .64787    .60422    .49342    .53014    .91246    .83299   1.00000
AVERAGE CORRELATION IS     .5107049264     TRACE OF FACTOR ANALYSIS MATRIX IS     9.000000000
S.E. OF A LOADING TIMES SQUARE ROOT OF THE NO. OF OBSERVATIONS IS     .7687631267

COMMUNALITY ESTIMATES(IE DIAGONAL OF MATRIX TO BE FACTOR ANALYZED)
  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000  1.00000

COMPUTED NUMBER OF FACTORS =  4  RANK OF CORRELATION INVERSE =   9
                          0/0 OF NUMBER    ACCUMULATIVE    0/0 OF COMMON      ACCUMULATIVE 0/0
 I    EIGENVALUES         OF VARIABLES     PERCENTAGE      VARIANCE(TRACE)    COMMON VARIANCE

 1     5.155875             57.29            57.29             57.29             57.29
 2     2.235041             24.83            82.12             24.83             82.12
 3     1.067853             11.87            93.99             11.87             93.99
 4      .239166              2.66            96.64              2.66             96.64
```

continued

```
        PRINCIPAL FACTOR LOADING MATRIX   WITH   4 FACTORS AND      0 CYCLES    IMAGE COR.EIG RANGE
              1    .961   -.749   -.532   -.340    .031
              2    .960   -.770   -.513   -.317    .058
              3    .977   -.833   -.351   -.392   -.081
              4    .955   -.702    .675   -.078   -.029
              5    .954   -.573    .753   -.165    .178
              6    .962   -.639    .695   -.184   -.192
              7    .977   -.802   -.119    .509    .243
              8    .981   -.752   -.217    .525   -.305
              9    .971   -.932   -.005    .310    .080
VARIANCE     10   8.698   5.156   2.235   1.068    .239
0/0TRACE     10  96.644  57.287  24.834  11.865   2.657

                    .               .              .
                    .               .              .
                    .               .              .

          VARIMAX FACTOR LOADING MATRIX   WITH   4 FACTORS AND      2 CYCLES    IMAGE COR.EIG RANGE
              1    .961   -.944    .027    .261    .016
              2    .960   -.934    .047    .289    .043
              3    .977   -.925    .232    .241   -.092
              4    .955   -.108    .936    .257   -.022
              5    .954   -.037    .952    .103    .188
              6    .962   -.116    .948    .131   -.184
              7    .977   -.270    .192    .903    .228
              8    .981   -.277    .089    .891   -.321
              9    .971   -.399    .401    .804    .068
VARIANCE     10   8.698   2.956   2.944   2.557    .240
0/0TRACE     10  96.644  32.841  32.715  28.416   2.672

SORTED    VARIMAX FACTOR LOADING MATRIX   WITH   4 FACTORS AND      2 CYCLES    IMAGE COR.EIG RANGE
      5  -.037     5   .952     7   .903     7   .228
      4  -.108     6   .948     8   .891     5   .188
      6  -.116     4   .936     9   .804     9   .068
      7  -.270     9   .401     2   .289     2   .043
      8  -.277     3   .232     1   .261     1   .016
      9  -.399     7   .192     4   .257     4  -.022
      3  -.925     8   .089     3   .241     3  -.092
      2  -.934     2   .047     6   .131     6  -.184
      1  -.944     1   .027     5   .103     8  -.321
     10  2.956    10  2.944    10  2.557    10   .240
     10 32.841    10 32.715    10 28.416    10  2.672
```

continued

```
FACTOR ANALYSIS                                        INDEP SCALE.BINARY.
TOTAL VARIABLES=   9. INPUT OPTION= -1. INTERMEDIATE OUTPUT=  2. TRANSFORM=-0. NSMC= 6. SELECT RANGE=-0
LIMITS--EIGENVALUE= 1.00. ROTATION CYCLES=  40. ROTATION ANGLE IN RADIANS= .005
PRINCIPAL-FACTOR= 1    QUARTIMAX= 5    -0.  -0.  1    VARIMAX= 7    -0.  -0.  1
EXTRA SPACE IN X ARRAY FOR TRANSGENERATION =     0 MAXIMUM NUMBER OF FACTORS =   3
INPUT DATA FORMAT  (9F4.3)
$$$ SAVE CORE SPACE. UMST550 CAN USE 070100B LESS FIELD LENGTH ON JOB CARD TO RUN THIS PROBLEM

 CORRELATION MATRIX                                    INDEP SCALE.BINARY.

     1    1.00000
     2     .82900    1.00000
     3     .76800     .77500    1.00000
     4     .10800     .11500     .27200    1.00000
     5     .03300     .06100     .20500     .63600    1.00000
     6     .10800     .12500     .23800     .62600     .70900    1.00000
     7     .29800     .32300     .29600     .24900     .13800     .19000    1.00000
     8     .30900     .34700     .27100     .18300     .09100     .10300     .65400    1.00000
     9     .35100     .36900     .38500     .36900     .25400     .29100     .52700     .54100    1.00000
INDEPENDENT SCALE FACTOR ANALYSIS

 IMAGE OR INDEPENDENT SCALE MATRIX                             INDEP SCALE.BINARY.

     1    2.74630
     2    2.63378    2.91161
     3    2.30359    2.36759    2.23132
     4     .33334     .41662     .55193    1.04652
     5     .19303     .20302     .46532    1.06613    1.33651
     6     .29572     .34886     .64926    1.06866    1.12265    1.28532
     7     .78880     .86815     .71765     .43315     .34178     .33620     .92609
     8     .82511     .85699     .77388     .36060     .17581     .30256     .77720     .99577
     9     .86470     .94177     .85065     .54331     .50139     .52829     .77183     .73063     .77261
AVERAGE CORRELATION IS     .7586096450     TRACE OF FACTOR ANALYSIS MATRIX IS     14.25203969
S.E. OF A LOADING TIMES SQUARE ROOT OF THE NO. OF OBSERVATIONS IS     .3404075418

COMMUNALITY ESTIMATES(IE DIAGONAL OF MATRIX TO BE FACTOR ANALYZED)
  2.74630   2.91161   2.23132   1.04652   1.33651   1.28532    .92609    .99577    .77261


COMPUTED NUMBER OF FACTORS =  3  RANK OF CORRELATION INVERSE =   9

                      0/0 OF NUMBER    ACCUMULATIVE    0/0 OF COMMON     ACCUMULATIVE 0/0
  I   EIGENVALUES     OF VARIABLES     PERCENTAGE      VARIANCE(TRACE)   COMMON VARIANCE

  1    8.956910          99.52           99.52            62.85            62.85
  2    3.295564          36.62          136.14            23.12            85.97
  3    1.238341          13.76          149.90             8.69            94.66
```

continued

```
PRINCIPAL FACTOR LOADING MATRIX   WITH   3 FACTORS AND     0 CYCLES    INDEP SCALE.BINARY+
              1   2.639  -1.540   -.486   -.181

              2   2.797  -1.602   -.457   -.145

              3   2.156  -1.437   -.144   -.264

              4    .992   -.481    .869   -.074

              5   1.235   -.398   1.011   -.232

              6   1.193   -.481    .956   -.217

              7    .844   -.646    .199    .622

              8    .893   -.648    .107    .679

              9    .741   -.703    .295    .401

VARIANCE     10  13.491   8.957   3.296   1.238

0/0TRACE     10  94.659  62.847  23.123   8.689


QUARTIMAX FACTOR LOADING MATRIX   WITH   3 FACTORS AND     3 CYCLES    INDEP SCALE.BINARY+
              1   2.639  -1.624   -.048    .012

              2   2.797  -1.671   -.013    .063

              3   2.156  -1.443    .268   -.031

              4    .992   -.232    .957    .152

              5   1.235   -.134   1.103    .007

              6   1.193   -.226   1.068    .028

              7    .844   -.485    .221    .748

              8    .893   -.505    .123    .789

              9    .741   -.540    .370    .559

VARIANCE     10  13.491   8.418   3.549   1.524

0/0TRACE     10  94.659  59.064  24.903  10.692


QUARTIMAX FACTOR SCORING MATRIX  3 FACTORS.   9 VARIABLES  INDEP SCALE.BINARY+
      1 1   -.222   -.068   -.136
      2 1   -.223   -.064   -.105
      3 1   -.192    .041   -.187
      4 1    .013    .275   -.006
      5 1    .018    .340   -.125
      6 1    .006    .324   -.114
      7 1    .010   -.026    .510
      8 1    .008   -.062    .550
      9 1   -.010    .039    .342
```

continued

```
         VARIMAX FACTOR LOADING MATRIX   WITH   3 FACTORS AND      7 CYCLES    INDEP SCALE.BINARY+
            1   2.639  -1.566    .044    .429
            2   2.797  -1.596    .077    .493
            3   2.156  -1.379    .352    .361
            4    .992   -.113    .953    .266
            5   1.235   -.044   1.105    .109
            6   1.193   -.130   1.074    .151
            7    .844   -.260    .189    .860
            8    .893   -.276    .089    .899
            9    .741   -.351    .355    .701
VARIANCE   10  13.491   7.203   3.585   2.704
O/OTRACE   10  94.659  50.537  25.152  18.970

SORTED   VARIMAX FACTOR LOADING MATRIX   WITH   3 FACTORS AND      7 CYCLES    INDEP SCALE.BINARY+
      5  -.044      5  1.105     8   .899
      4  -.113      6  1.074     7   .860
      6  -.130      4   .953     9   .701
      7  -.260      9   .355     2   .493
      8  -.276      3   .352     1   .429
      9  -.351      7   .189     3   .361
      3 -1.379      8   .089     4   .266
      1 -1.566      2   .077     6   .151
      2 -1.596      1   .044     5   .109
     10  7.203     10  3.585    10  2.704
     10 50.537     10 25.152    10 18.970

  VARIMAX FACTOR SCORING MATRIX  3 FACTORS,   9 VARIABLES  INDEP SCALE,BINARY+
       1 1  -.254  -.044  -.078
       2 1  -.247  -.043  -.047
       3 1  -.229   .066  -.128
       4 1   .032   .274   .008
       5 1   .011   .347  -.104
       6 1   .001   .331  -.091
       7 1   .138  -.067   .487
       8 1   .143  -.106   .525
       9 1   .080   .012   .335
```

COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 2.007 seconds. At the current rate for the Univ. of Minnesota ($0.20/sec.), the computer time cost $0.40 plus a charge for output supplies.

Charge to user = computer costs + postage + network overhead
= $1.52 + postage + network overhead

CONTENTS—UMST550

pages	
1– 4	Identification & Abstract
5–18	User Instructions
19–32	I/O
33	Cost—Contents

DESCRIPTIVE TITLE	Least Square Curve Fitting Using Orthogonal Polynomials
CALLING NAME	UMST560
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

Using a modified method of least squares, this program fits polynomials to pairs of input data vectors. It first calculates a polynomial of degree 1, then of degree 2, and so on up to the degree of the polynomial specified.

Output and Options

Different weights for each data element may be specified.
After every iteration the error sum of squares and standard error of the coefficients are printed.
After every iteration observed values for both X and Y, predicted Y, observed minus predicted Y, and weights assigned to each point may be requested. In this notation X is the independent variable and Y the dependent variable. This option will be called the Full Output. It is useful information when the degree of the polynomial desired is known.
An automatic termination with the polynomial of best fit may be selected.
The coefficients of the orthogonal polynomials used in the calculation of the least squares polynomials may be printed.
The input data may be transformed and vectors added according to specifications supplied to the program. Pairs of X—Y vectors are selected for curve fitting if there are more than two variables. By use of transformation 04,

continued

$x' = e^x$, an exponential fit of the data may be obtained of the form $y = a_0 + a_1 e^x + a_2 e^{2x} + \ldots$

A graph of the polynomial together with observed values may be printed.

Method

Let X_i denote the independent variable (abscissa), Y_i the dependent variable (ordinate), W_i the corresponding weight ($W_i = 1$ if weights are not supplied), and $P_N(X) = \sum_{i=0}^{N} B_i X^i$ the least-square polynomial of degree N. If $P_J(X)$ denotes any Jth-degree polynomial, then $P_N(X)$ is that polynomial which minimizes $\sum_{i=1}^{M} W_i [Y_i - P_J(X_i)]^2$.

Because of the much greater accuracy attainable, orthogonal polynomials, rather than the standard matrix inversion technique, are used to determine $P_N(X)$. There is no restriction on the spacing of the X_i. Calculation and printout begin with degree 1 and proceed to the highest degree selected.

If the automatic termination option is selected, the program terminates when the "error sum of squares", a standard criterion of computational accuracy, is not reduced for two consecutive higher degrees. The full output and differences are then printed out for that polynomial which resulted in the smallest "error sum of squares". This process constructively defines the polynomial of "best fit" since the probability is small that polynomials of higher degree will yield a smaller "error sum of squares."

Equations

$M_W = \sum_{i=1}^{M} W_i$; M = number of data points

M_N = maximum of (1, M–N–1) ; N = degree of polynomial

$P_N(X) = \sum_{i=1}^{M} B_i X^i$; least squares polynomial

continued

Means:

a. $\bar{X} = \sum_{i=1}^{M} W_i X_i / M_W$

b. $\bar{Y} = \sum_{i=1}^{M} W_i Y_i / M_W$

Total sum of squares:

$$Y_s = \sum_{i=1}^{M} W_i Y_i^{2} - \bar{Y}^2 \cdot M_W$$

Standard deviations:

a. $S_X = [\sum_{i=1}^{M} W_i X_i^{2} - \bar{X}^2 \cdot M_W] / (M_W - 1)$

b. $S_Y = Y_S / (M_W - 1)$

Error sum of squares:

$$E = \sum_{i=1}^{M} W_i [P_N(X_i) - Y_i]^2$$

Standard Error of Y Estimate:

$$S_E = \sqrt{E/M_N}$$

Note: S is undefined when N = M–1. In this case $M_N = 1$ is used in the calculation.

Standard Error of the Coefficients:

$SB_i = \sqrt{(\text{Variance } B_i)/M_N}$ See Ref. 1 for computation of these. See also previous note on M_N.

Back solution:

$P_N(X_i)$ = Predicted Y ; Observed Y – Predicted Y = $Y_i - P_N(X_i)$

continued

Time Estimate

The maximum time a program will run is a function of the number of observations, the number of variables, the number of Select Cards, and the number of graphs printed. The equation given below assumes a graph size of 50 by 100 lines and a Full Output option.

$$\text{TIME} = [.15\ \text{NVAR} + \text{NSELE} * \text{NOBS}^{.38} + \text{NOBS}\ (.00036) + 4.62 * \text{NGRAF}]\ \text{Seconds}$$

Where: NSELE = the number of Select Cards
NVAR = the number of variables
NOBS = the number of observations
NGRAF = the number of 50 X 100 graphs printed out.

REFERENCES

1. Todd, J., *A Survey of Numerical Analysis* (New York: McGraw-Hill Book Co., Inc., 1962), pp. 347–362.

2. Hildebrand, F.B., *Introduction to Numerical Analysis* (New York: McGraw-Hill Book Co., Inc., 1956) pp. 258–287.

3. Jordan, T.L., and Vogel, R.E., "Least Squares Polynomial Fitting," *Systems Programs for CO-OP*, E2 UCSD LSQPOL, (Palo Alto, Calif: Control Data Corp.).

4. Hamming, R.W., *Numerical Methods for Scientists and Engineers* (New York: McGraw-Hill Book Co., Inc. 1962).

USER INSTRUCTIONS

System Control Cards

System Control Cards will be prepared by University of Minnesota personnel.

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7	ID	PROBLEM
8–11	M	Number of data points. $M \geq 1$ *Warning:* It is recommended that many data points be used to get a better curve fit. A suggested value is 40*N. K data points will fit a K–1 degree polynomial exactly. That is, if only K points are used the error between the calculated curve and the data points is zero, but it is unlikely that such a curve would fit data points between the given ones so such a small number of points should be avoided.
12–14	NVAR	Number of variables. $NVAR \geq 2$
15–17	N	The highest degree polynomial wanted, $1 \leq N \leq M-1$ and $N \leq 30$.
18	NFC	Number of Format Cards. $NFC \leq 9$ If NFC is zero or blank, one Format Card is assumed.
19	W	1: weights are used 0: otherwise
20	ISW	1: orthogonal polynomials are to be printed 0: otherwise
21	IP	1: Full Output desired for polynomials of degree 1 to degree N. 0: only intermediate statistics and final Full Output will be printed.

continued

Columns	*Parameter*	*Contents*
22	IA	1: curve fitting should cease when the 'best fit' is found. (See section on Method in Functional Abstract for definition of best fit.)
23–25	XNTX	The number of spaces the X coordinate of the graph will occupy. XNTX $\geq$ 1 (See section on graphing option.)
26–28	YNTY	The number of spaces the Y coordinate of the graph will occupy. 1 $\leq$ YNTY $\leq$ 125
29	MODG	0: the graph is not desired 1: graph desired with each polynomial fit 2: graph desired only on final output
30	MODREAD	0: data order is XYXY, or WXYWXY if weights are used; $X_1X_2X_3...X_1X_2X_3$ ... or $WX_1X_2X_3...WX_1X_2X_3...$ if more than 2 variables are used 1: data order is YXYX, etc. or WYXWYX if weights are used. This option must be used only when NTRNGEN is 0 and when NVAR is 2.
31–33	NTRNGEN	The number of transformations. See NTRANS Cards below. NTRNGEN $\geq$ 0
34–36	KNEW	The number of new vectors added by the transformation. See NTRANS Cards below. KNEW $\geq$ 0
37–80	NAME	The alphanumeric identification of the problem.

Format Cards

The number of Format Cards is given by NFC on the Problem Card. If NFC is zero or blank, one Format Card is assumed. The Format Card(s) must provide for reading all elements in a data set using X, F, or E fields. If NTRNGEN is zero and NVAR is two, a data set consists of all M data points in reverse order (and weights if used) and thus the MODREAD option 1 is assumed (see above). If NTRNGEN is not zero and NVAR > 2, a data set consists of all variables associated with a data point.

continued

Input Data Cards

Data Cards may be punched in one of three ways:

a. NTRNGEN $\geq$ 0 and NVAR > 2. Denote the first variable by X_1, the second variable by X_2 and so on. Then punch the cards in the form $WX_1X_2...X_m$ if weights are used, or $X_1X_2...X_m$ if weights are not used.

b. NTRNGEN = 0 and NVAR = 2 and MODREAD = 0, then let X_i denote the independent variable (abscissa of the data point), Y_i denote the dependent variable (ordinate of the data point), and W_i the corresponding weight if used. The Format Card shall provide for reading the data in the form: $W_1X_1Y_1W_2X_2Y_2...W_mX_mY_m$ or $X_1Y_1X_2Y_2...X_mY_m$

c. NTRNGEN = 0 and NVAR = 2 and MODREAD = 1. The Format Card shall provide for reading the data in the form: $W_1Y_1X_1W_2Y_2X_2...W_mY_mX_m$ or $Y_1X_1Y_2X_2...Y_mX_m$

NTRANS Card(s)

Description of Transgernation

The total number of NTRANS Cards is specified by NTRNGEN on the Problem Card. Each NTRANS Card specifies the transformation or variable generation desired. The new variable or transformed variable may replace one of the existing variables or may be added to the original set of variables. Transformations are done sequentially; thus the sum of two variables may be generated by one transformation and the square root of the sum may be done by the next transformation. All variables added by the transgeneration have weights equal to one. The total size of the array saved by the program is NVAR + KNEW. Thus, do not include in KNEW the transgenerations which replace other variables. Violation of restrictions relating to the transformations, transgenerations, and maximum value of NVAR + KNEW result in error messages, and after all the cards have been checked the problem is aborted.

Notation

x: the variable prior to transgeneration

x': the transformed variable which (depending on instructions in the NTRANS Card) replaces the original variable x or becomes a new variable

c: a constant

A,B: specified variable numbers, or indices

*: multiplication

continued

Code	*Transformation*	*Restriction*
00	$x' = x$	
01	$x' = \sqrt{x}$	$x \geq 0$
02	$x' = \sqrt{x} + \sqrt{x+1}$	$x \geq 0$
03	$x' = \log_{10} x$	$x \geq 0$
04	$x' = e^{x}$	$x \leq 740$
05	$x' = \sin^{-1} x$	$-1 \leq x \leq 1$
06	$x' = \log_{e} x$	$x > 0$
07	$x' = \sin(x)$	$\lvert x \rvert \leq 2.2 \times 10^{14}$, x in radians
08	$x' = \cos(x)$	$\lvert x \rvert \leq 2.2 \times 10^{14}$, x in radians
09	$x' = \frac{1}{x}$	$x \neq 0$
10	$x = x + c$	
11	$x' = x*c$	
12	$x' = x^{c}$	$x > 0$
13	$x' = c^{x}$	$c > 0$
14	If $x \geq c$ set $x' = 1$ If $x < c$ set $x' = 0$	
15	If $x_a \geq x_b$ set $x' = 1$ If $x_a < x_b$ set $x' = 0$	
16	$x' = x_a + x_b$	
17	$x' = x_a - x_b$	
18	$x' = x_a * x_b$	
19	$x' = x_a/x_b$	$x_b \neq 0$
20	$x' = x_a^{x_b}$	$x_a > 0$
21	If $x' = C_1$ set $x' = 1$ If $x' = C_2$ set $x' = -1$ Otherwise set $x' = 0$	

continued

Code	*Transformation*	*Restriction*
21	(Notice that the x' variable (no. indicated in Col. 7–11) is replaced by 1, –1, or 0. If you do not want to destroy your original variable, it is necessary to create a new one by using transformation 00, then using transformation 21 on the new variable.)	

NTRANS Cards

Columns	*Parameter*	*Contents*
1– 6		NTRANS
7–11	x'	Variable number
12–16	LL	Transgeneration code
17–21		Variable number of x, x_a, or constant C_1 for transformation 21
22–31		Variable number of x_b, constant c, or constant C_2 for transformation 21 (keypunch decimal place if not at extreme right)

Select Card

This card gives the location of the independent and dependent variables in the array. As many Select Cards as desired may be included in any problem. If MODREAD option is used (NTRNGEN = 0, NVAR = 2) no Select Card is necessary.

Columns	*Parameter*	*Contents*
1– 6		SELECT
11–13	IX	The no. of the independent variable
14–16	IY	The no. of the dependent variable

Finish Card

Columns	*Content*
1– 6	FINISH

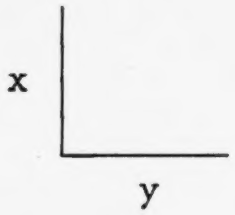
A Note on the Graphing Option

The graphing program orders the data on the independent variable, then transforms the data to the dimensions specified on the

continued

Problem Card. It will allow more than one Y value for a given X value; thus, for example, the two points (5,3) and (5,7) will both appear on the graph. The abscissa (X axis) is the vertical axis and the ordinate (Y axis) is the horizontal axis, thus turn the output 90 degrees counter clockwise to obtain a correct picture of the polynomial. See Sample I/O for an example of the output.

Orientation



Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

Problem Card
Format Card(s)
Data Cards
NTRANS Card(s) if necessary
Select Card(s) if necessary
repeat for additional problems
Finish Card

SAMPLE INPUT

```
PROBLEM   5  3  310100  0  000  3  3UMST560 TEST CHECK TRNSGEN
(3F5.0)
    1    2    2
    5    7    6
    6    8    7
    2    3    3
    6    5    7
NTRANS    1     0     1
NTRANS    4    12     1          3
NTRANS    5     9     1
SELECT       1  5
SELECT       1  2
SELECT       1  4
PROBLEM   5  2  310110 501002000000 MST 56 TEST DECK SIMPLE CASE NEGATIVE VALUES
(15F4.0)
 001-200 006 001-100 003 001 200  -6 001 100  -3 001 800 -24
FINISH
```

```
POLYNOMIAL CURVE FIT OF DEGREE  1


ERROR SUM OF SQUARES =     7.240404040E-02                    STD. ERROR OF Y EST. =      1.553534040E-01


LEAST SQUARE COEFFICIENTS. B(0), B(1), B(2), ETC.
    9.763636364E-01    -1.424242424E-01

STANDARD ERROR OF THE COEFFICIENTS, SB(0), SB(1), SB(2), ETC.
    1.495975624E-01     3.312145702E-02




POLYNOMIAL CURVE FIT OF DEGREE  2


ERROR SUM OF SQUARES =     1.574803150E-02                    STD. ERROR OF Y EST. =      8.873565094E-02


LEAST SQUARE COEFFICIENTS. B(0), B(1), B(2), ETC.
    1.437532808E+00    -5.401574803E-01     5.538057743E-02

STANDARD ERROR OF THE COEFFICIENTS, SB(0), SB(1), SB(2), ETC.
    1.919871707E-01     1.494767363E-01     2.064584347E-02




POLYNOMIAL CURVE FIT OF DEGREE  3


ERROR SUM OF SQUARES =     1.937442383E-27                    STD. ERROR OF Y EST. =      4.401638767E-14


LEAST SQUARE COEFFICIENTS. B(0), B(1), B(2), ETC.
    1.866666667E+00    -1.083333333E+00     2.333333333E-01    -1.666666667E-02

STANDARD ERROR OF THE COEFFICIENTS, SB(0), SB(1), SB(2), ETC.
    1.781168690E-13     2.044400766E-13     6.325202385E-14     5.845881535E-15
```

continued

```
I        X(I)               OBSERVED Y          PREDICTED Y        OBS. Y - PRED. Y        WEIGHT

1   1.000000000E+00    1.000000000E+00    1.000000000E+00     7.105427358E-15    1.000000000E+00
2   5.000000000E+00    2.000000000E-01    2.000000000E-01    -1.776356839E-14    1.000000000E+00
3   6.000000000E+00    1.666666667E-01    1.666666667E-01    -1.953992523E-14    1.000000000E+00
4   2.000000000E+00    5.000000000E-01    5.000000000E-01    -2.842170943E-14    1.000000000E+00
5   6.000000000E+00    1.666666667E-01    1.666666667E-01    -1.953992523E-14    1.000000000E+00


ORTHOGONAL POLYNOMIAL COEF FOR K =     2

-4.0000000E+00  1.0000000E+00


ORTHOGONAL POLYNOMIAL COEF FOR K =     3

 8.3272727E+00 -7.1818182E+00  1.0000000E+00


ORTHOGONAL POLYNOMIAL COEF FOR K =     4

-2.5748031E+01  3.2590551E+01 -1.0677165E+01  1.0000000E+00



              .                      .                      .
              .                      .                      .
              .                      .                      .
```

continued

```
UMST560   LEAST SQUARE CURVE FITTING USING ORTHOGONAL POLONOMIALS   MST 56 TEST DECK SIMPLE CASE NEGATIVE VALUES


  DATA POINTS      5      NUMBER OF VARIABLES     2      MAXIMUM DEGREE     3      FORMAT CARDS    1

  WEIGHTS    0 =  NO.      ORTHO. POLY. PRINTED    1 = YES.      FULL OUTPUT    1 = YES.      STOP AT BEST FIT    0 =  NO.

  NUMBER OF X SPACES    50      NUMBER OF Y SPACES   100      GRAPH    2 = LAST POLY.

  NUMBER OF NTRANS CARDS      0   VECTORS ADDED BY TRANSFORMATION      0

  DATA ORDER (ONLY VALID IF NUMBER OF NTRANS CARDS = 0 AND NUMBER OF VARIABLES = 2)    0 = XY OR WXY


  FORMAT = (15F4.0)


        THIS PROBLEM REQUIRES A FIELD LENGTH OF ONLY 032500B



                                              MEAN              STD. DEVIATION

                                    X       -1.960000000E+01        4.514753592E+01

                                    Y        1.000000000E+00        1.414266594E+02

                                       TOTAL SUM OF SQUARES =  8.000600000E+04


  POLYNOMIAL CURVE FIT OF DEGREE  1


  ERROR SUM OF SQUARES =      8.000099534E+04                     STD. ERROR OF Y EST. =      1.633003320E+02


  LEAST SQUARE COEFFICIENTS. B(0), B(1), B(2), ETC.
     5.143992543E-01    -2.477554825E-02

  STANDARD ERROR OF THE COEFFICIENTS. SB(0), SB(1), SB(2), ETC.
     8.117812032E+01     1.808518767E+00


   I        X(I)             OBSERVED Y         PREDICTED Y       OBS. Y - PRED. Y        WEIGHT

   1    1.000000000E+00    -2.000000000E+02    4.896237060E-01    -2.004896237E+02    1.000000000E+00
   2    6.000000000E+00     1.000000000E+00    3.657459648E-01     6.342540352E-01    1.000000000E+00
   3   -1.000000000E+02     3.000000000E+00    2.991954079E+00     8.045920620E-03    1.000000000E+00
   4    1.000000000E+00     2.000000000E+02    4.896237060E-01     1.995103763E+02    1.000000000E+00
   5   -6.000000000E+00     1.000000000E+00    6.630525438E-01     3.369474562E-01    1.000000000E+00
```

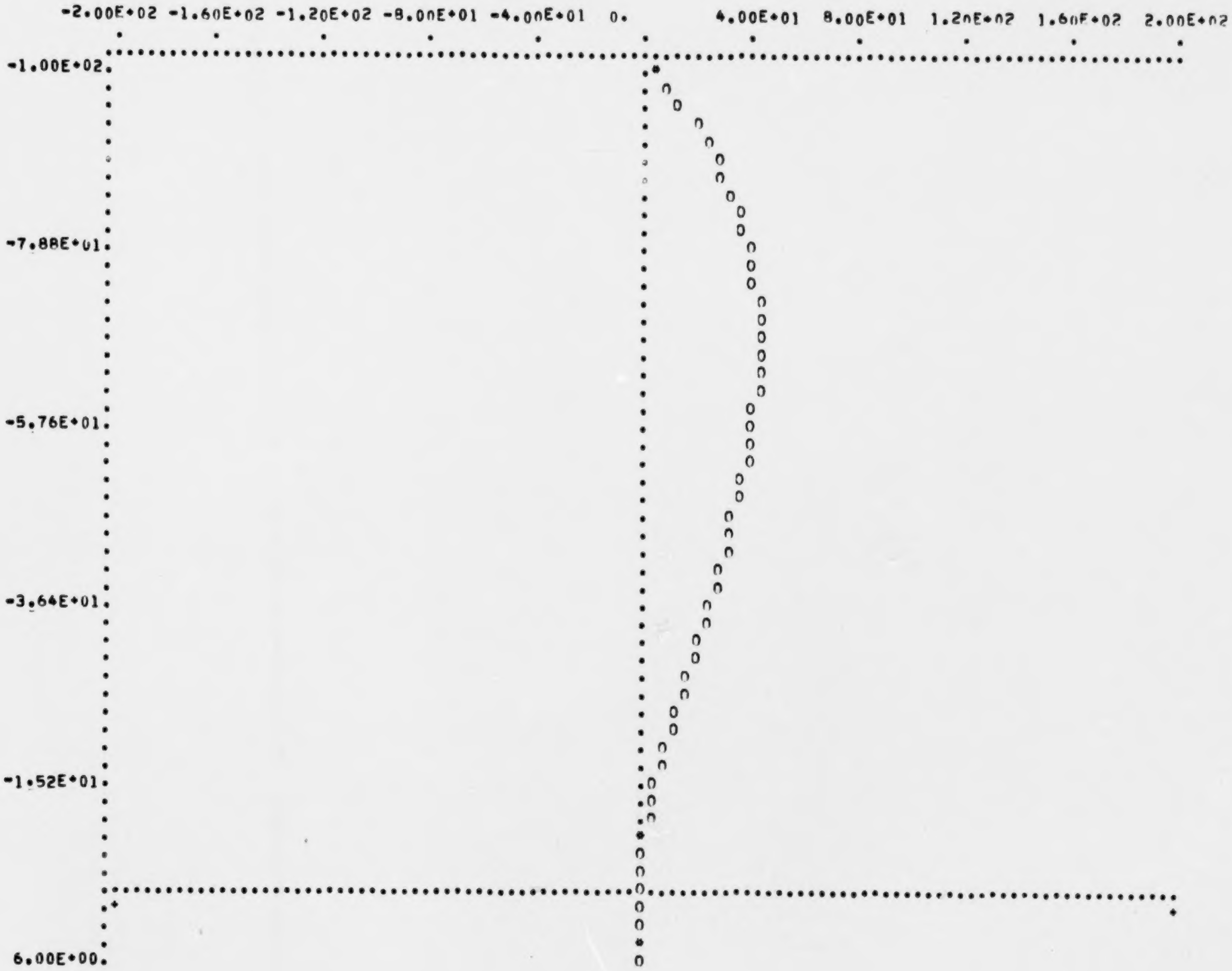
continued

```
POLYNOMIAL CURVE FIT OF DEGREE  2


ERROR SUM OF SQUARES =     8.000097684E+04                      STD. ERROR OF Y EST. =      2.000012210E+02


LEAST SQUARE COEFFICIENTS. B(0), B(1), B(2), ETC.

    5.005774899E-01    -8.863804601E-03     1.616497617E-04


STANDARD ERROR OF THE COEFFICIENTS. SB(0), SB(1), SB(2), ETC.

    1.014785784E+02     2.350202346E+01     2.376977996E-01


  I         X(I)              OBSERVED Y          PREDICTED Y       OBS. Y - PRED. Y       WEIGHT

  1    1.000000000E+00    -2.000000000E+02     4.918753350E-01    -2.004918753E+02    1.000000000E+00
  2    6.000000000E+00     1.000000000E+00     4.532140537E-01     5.467859463E-01    1.000000000E+00
  3   -1.000000000E+02     3.000000000E+00     3.003455567E+00    -3.455567388E-03    1.000000000E+00
  4    1.000000000E+00     2.000000000E+02     4.918753350E-01     1.995081247E+02    1.000000000E+00
  5   -6.000000000E+00     1.000000000E+00     5.595797089E-01     4.404202911E-01    1.000000000E+00


POLYNOMIAL CURVE FIT OF DEGREE  3


ERROR SUM OF SQUARES =     8.000000000E+04                      STD. ERROR OF Y EST. =      2.628427125E+02


LEAST SQUARE COEFFICIENTS. B(0), B(1), B(2), ETC.

   -1.845909773E-02    -1.011233084E-02     2.829053049E-02     2.808980789E-04


STANDARD ERROR OF THE COEFFICIENTS, SB(0), SB(1), SB(2), ETC.

    2.065392402E+02     3.323859785E+01     8.056829411E+00     8.038632014E-02


  I         X(I)              OBSERVED Y          PREDICTED Y       OBS. Y - PRED. Y       WEIGHT

  1    1.000000000E+00    -2.000000000E+02     9.725553696E-13    -2.000000000E+02    1.000000000E+00
  2    6.000000000E+00     1.000000000E+00     1.000000000E+00     1.747935130E-12    1.000000000E+00
  3   -1.000000000E+02     3.000000000E+00     3.000000000E+00     2.188471626E-12    1.000000000E+00
  4    1.000000000E+00     2.000000000E+02     9.725553696E-13     2.000000000E+02    1.000000000E+00
  5   -6.000000000E+00     1.000000000E+00     1.000000000E+00     2.330580173E-12    1.000000000E+00


ORTHOGONAL POLYNOMIAL COEF FOR K =      2

 1.9600000E+01  1.0000000E+00


ORTHOGONAL POLYNOMIAL COEF FOR K =      3

 -8.5504391E+01  9.8433449E+01  1.0000000E+00
```

continued

```
ORTHOGONAL POLYNOMIAL COEF FOR K =     4
-1.8477755E+03 -4.4447660E+00  1.0013910E+02  1.0000000E+00
```

```
                        GRAPH OF VARIABLE   1 VERSUS VARIABLE   2
                               LEGEND
                                +   OBSERVED POINT
                                O   CALCULATED POINT
                                *   OBSERVED AND CALCULATED POINT COINCIDE
                                W   CALCULATED VALUE OUTSIDE THE RANGE OF THE GRAPH
      VARIABLE    2 IS  PLOTTED ALONG THIS AXIS
```



COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 0.755 seconds. At the current rate for the University of Minnesota ($0.20/sec.), the computer time cost $0.15 plus a charge for output supplies.

Charge to user = computer time + postage + network overhead
= $0.53 + postage + network overhead

CONTENTS—UMST560

pages	
1– 4	Identification & Abstract
5–10	User Instructions
11–17	I/O
19	Cost—Contents

635A

DESCRIPTIVE TITLE	Multivariate Analysis of Variance
CALLING NAME	UMST570
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	William Walster, David Doren University Computer Center University of Minnesota
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program performs multivariate analysis of variance and/or covariance on a vector of dependent variables.

Input

The input data can be from one to many dependent variates and/or covariates. Each observation *must* have a complete set of scores on all the variates and covariates. It may have unequal observations in the cells of the selected design.

Options

Linear and/or eleven algebraic transformations can be made when data are read in.

The analyses of all anova designs are based on contrasts (comparisons). The program generates several main effect options for between analyses or allows freedom for the programmer to supply his own design matrix (set of contrasts). The program forces orthogonality among contrasts. All the above contrasts are for "between-cells analyses".

continued

Sources of Variation

a. The program automatically groups the appropriate sets of contrast estimates into main effects and appropriate interaction effects. Options are available to regroup the contrast estimates. This allows tests on specific contrasts or contrast sets other than the normal main effects and interaction effects.

b. *The order of the contrasts in the design matrix are of a specific form.* It is advised that the user acquaint himself with the section on Regrouping Cards to fully understand the "conventional" order utilized by the program. This "conventional" order may be changed by the user through the "reorder" option (see section on Special Orderings). The order of the tests of the sets of contrasts takes on special meaning since each test assumes the previous test was null.

A test of the equality of the variance-covariance matrices between cells can be performed. When there is only one dependent variate, this test is the same as the test of equality of variances.

Either weighted or unweighted means analysis may be chosen independently for each factor in the design.

A second linear transformation of variates can be performed to allow orthogonal linear combinations of original variates within each observation. The resulting analysis, utilizing proper recombinations, can give a multivariate analog to univariate repeated-measures analyses.

The recombination option allows the design analysis to be performed on selected subsets of variates (and covariates).

USER INSTRUCTIONS

System Control Cards

System control cards will be prepared by University of Minnesota personnel.

Program Control Cards

Problem Card (Required)

Format (A7, 5I2, 2I3, 21I2) It is *strongly* suggested that the user *completely* lay out his design and the options to be used prior to completing the Problem Card.

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8– 9	N	Problem number (01 through 98)
10–11	IVAR	Number of original variates for analysis of variance
12–13	ICOVAR	Number of original covariates (concomitant) variables for analysis of covariance
14–15	NFACT	Number of factors in the design
16–17	MFACT	Number of levels in the largest factor of the design
18–20	NCELLS	The number of cells in the design
21–23	NTESTS	Number of sources of variation to be tested. Ordinarily includes interactions, main effects, and almost always includes the "grand mean."
24–25	KNDEX	Blank: number "1" through "K" are to be used as indices of original variates and covariates where the number of variates plus covariates equals K 01: indices of original variates and covariates are to be supplied. (See section on Special Index Cards for original variates.)

continued

Columns	*Parameter*	*Contents*
26–27	KNDEX1	Blank: numbers "1" through "K" are to be used as indices of new variates and covariates after linear transformation No. 1, where the number of new variates plus covariates is equal to K. 01: indices are to be supplied. (Col. 40 and 41 must be 01 if this option is used; see section on Special Index Cards for new variates).
28–29	KNDEX2	Blank: number "1" through "K" are to be used as indices of new variates and covariates after linear transformation No. 2 where the number of new variates and covariates after transformation No. 2 equals "K". 01: indices are to be supplied. (Col. 46–47 must be 01 if this option is used; see section on Special Index Cards for new variates.)
30–31	1TRANS	Blank: no algebraic transformations are used. K: *K* algebraic transformations are desired. It should be noted at this point that the algebraic transformations are made after linear transformation No. 1 and before linear transformation No. 2 so that if linear transformation No. 1 is used then the algebraic transformations are made on the new variates and/or covariates after linear transformation No. 1. Any number of algebraic transformations may be used. See sections on Algebraic Transformation Card and Constant Card.
32–33	MORE	Blank: the within-cells sum of squares is to be used as the error term. 01: other sources of variation are to be added to the within-cells sum of squares to form the error term. See section on Augmenting.

continued

Columns	*Parameter*	*Contents*
34–35	1GROUP	Blank: the degrees of freedom are not to be regrouped. 01: the degrees of freedom are to be regrouped. See sections on Regrouping and Special Ordering.
36–37	KORDER	Blank: the statistical tests are to be calculated in the "conventional" order (the conventional order will be specified in section on Regrouping) 1: a reordering of the statistical tests is desired. If reordering is desired, the regrouping feature must be used to regroup degrees of freedom for tests; i.e. Cols. 28–29 must be 01. See section on Special Ordering.
38–39	NTS	Blank: data are read from input tape and not saved. 01: data are read from the data file and saved on scratch tape. 02: data are to be read from the scratch tape. This feature allows one to run consecutive problems on the same data without having duplicate data decks.
40–41	LT1	Blank: linear transformation No. 1 is not used. 01: linear transformation No. 1 is used. Linear transformation No. 1 is made on the data as they are read in. This option allows linear transformations on data before test of equality of variance-covariance matrices, and before algebraic transformations of the data are made. It should be noted at this point that when analysis of covariance is performed that the variates must come first, then the covariates. If the data are not arranged this way on cards one can use linear transformation No. 1 to reorder the variables in the computer. See section on Linear Transformation No. 1 Cards if 01.

continued

Columns	*Parameter*	*Contents*
42–43	NTVAR1	New number of *variates* after linear transformation No. 1
44–45	NTCOV1	New number of *covariates* after linear transformation No. 1
46–47	LT2	Blank: linear transformation No. 2 is not used 01: linear transformation No. 2 is used Linear transformation No. 2 is made *after* the test of equality of variance-covariance matrices and algebraic transformations. *Note:* This option is independent of linear transformation No. 1. See section on Linear Transformation No. 2 Cards.
48–49	NTVAR2	New number of *variates* after linear transformation No. 2
50–51	NTCOV2	New number of *covariates* after linear transformation No. 2
52–53	NADM	Blank: a test of equality of variance-covariance matrices is not desired 01: a test of equality of variance-covariance matrices is desired with the correlation coefficients for each cell *not* printed out 02: the test of equality of variance-covariance matrices is desired with the correlation matrices of each cell printed out. When there is only one dependent variable this test becomes a test of equality of variances.
54–55	NSKP	Blank: analysis of variance or multivariate analysis of variance should be performed on the variates and/or covariates that result after linear transformation No. 2 01: the program should skip to recombinations of variates and covariates. See section on Recombination Analysis Cards.

continued

Columns	*Parameter*	*Contents*
56–57	LNM	Blank: no check on orthogonalization is desired 01: a check on orthogonalization is desired This is a check on the orthogonalization of the design matrix. When the cell frequencies are unequal and the number of cells is large, rounding error can enter into the orthogonalization procedure and it is a good idea to check this. The program prints out a symmetric matrix which should be very close to the identity matrix. To the extent that it is not, rounding error has entered into the orthogonalization.
58–59	NTC	Blank: no Problem Title Cards are to be used K: *K* Problem Title Cards are used. See section on Problem Title Cards.
60–61	NXM	Blank: the cell means and marginal means are to be printed out. 01: the marginal means are not to be printed out 02: both the marginal means and the cell means are not to be printed out. This option is to save paper in case the same problem is being run a number of different ways.
62–63	NXA	Blank: the adjusted cell means and adjusted marginal means are to be printed out 01: the marginal adjusted means are not be be printed out 02: both the adjusted cell means and the adjusted marginal means are not to be printed out Again this option is to save paper.
64–65	NXD	01: the design matrix is to be omitted. This option is also to save paper.
66	NFC	Number of Format Cards. $0 \leq NFC \leq 9$. If zero or blank, NFC=1 is assumed.

continued

Problem Title Card(s) (Optional)

Format (10A8)
This option is included to aid in identifying different problem. NTC cards are read and 80 columns of each card will be read and printed out at the beginning of the problem output.

Contrast Option Card (Required)

Format (36I2)

Columns	*Parameter*	*Contents*
1– 2	NL1	Number of levels of factor 1. This is the slowest moving variable.
3– 4	IOP1	Contract option code of factor 1 (see below)
5– 6	NL2	Number of levels of factor 2
7– 8	IOP2	Contrast option code of factor 2
9–10	NL3	Number of levels of factor 3
11–12	IOP3	Contract option code of factor 3
⋮	⋮	⋮
71–72		

NFACT levels and contrast option codes are read, 18 pairs per card. Any number of factors may be used within storage limitations of the computer. An error message will be written out if too much storage is required for the particular problem.

Contrast Options

A number of options are available for setting up the design matrix. The codes and the corresponding types of contrast options are presented below. Each code is a two digit number. The first digit can either be a 0 or a 1. 0 indicates that the corresponding factor is to be weighted by the cell frequencies. If all of the factors are weighted then the analysis is a Least Squares Analysis. If the first digit of any factor is a 1 that indicates that that factor is to be unweighted. If all of the factors are unweighted the analysis would be normally called an Unweighted Means Analysis. It is possible with this program to have some factors weighted and other factors unweighted. This would be desirable in many designs where the experimenter had control over the numbers of observations in certain factors but not in others. The second digit of each contrast option code can take on the values 0 to 7.

continued

Code of Second Digit	*Type*
0	Nested Contrasts: This code indicates that all of more "minor" factors are to be nested in the given factor. *Note*: If there is one factor in which the rest of the factors are to be nested it has to be the first factor, factor 1, or the "major" factor. If there are two factors in which all of the rest of the factors are to be nested then those two have to be the first two factors, i.e., factors 1 and 2.
1	Nominal Contrasts: Where there are I levels in a factor, cell mean X_i is deviated from cell mean $\overline{X}_I$ for (i = 1, 2, ..., I–1).
2	Helmert Contrasts: Where there are I levels in a factor, cell mean X_i is deviated from the average of means X_{i+1} through X_I for (i = 1, 2, ..., I–1).
3	Optional one-way contrasts *on cell means* to be supplied. See section on Special One-Way Contrasts.
4	Optional one-way contrasts *on the parameters in the model* to be supplied. See section on Special One-Way Contrasts.
5	The optional design matrix to be used when Kronecker product of one-way design bases are not taken. See section on Special One-Way Contrasts.
6	Contrasts based on orthogonal polynomial coefficients. This option generates contrasts for testing linear, quadratic, ..., Ith–1 order effects. For this option it is assumed that there is an equal interval between all of the I treatment levels.
7	Contrasts based on orthogonal polynomial coefficients where there are not equal intervals. The interval widths must be specified in a card to follow. See section on Interval Width Cards.

Test Titles Cards (Required)

Format (3X,2A8)
This feature is included to ease the clerical burden of identifying results when they are printed out. The first three columns of each card are left blank so that an identifying code can be used to keep the cards in order. Sixteen characters are

continued

provided for the title in Cols. 4–20 of each card. The number of Title Cards must agree with Col. 21–23 of the Problem Card. When the Kronecker product scheme is used as programmed and no regrouping of the degrees of freedom is undertaken (see section on Regrouping) simply enter the title of each effect on a card with the cards in the "natural" order in which statistical tests are usually calculated. For the usual three-factor experiment, this will consist of the eight available tests in the "conventional" order: Grand Mean, C, B, BXC, A, AXC, AXB, AXBXC. AXBXC is then tested first.

Data Format Card (Required)

Format (10A8)
The Format Cards specify the format in F, E or X fields for the Data Cards. There must be NFC Format Cards (Col. 66 on Problem Card).

Special Index Card(s) for Original Variates and Covariates (Optional)

Format (18I4) Used if KNDEX = 01 (Cols. 24–25 on the Problem Card)
The standard indexing of variates is 1, 2, 3, ..., K where K is the sum of the number of variates and covariates, IVAR + ICOVAR (Cols. 10–11, Cols. 12–13 on the Problem Card). If special index numbers are desired in the output, the special indexing feature may be used.

Index numbers must be entered on the Special Index Card in the same order as the variables are entered on the Data Cards.

Columns	*Contents*
1– 4	Index of first variable
5– 8	Index of second variable
9–12	Index of third variable etc. in four-digit
⋮	fields through Col. 72.

Index numbers must be punched to the extreme right of each field. *Index number 9999 must not be used.* If there are more than 18 variables, punch as many cards as needed, 18 per card.

Special Index Card(s) for New Variables and Covariates after Linear Transformation No. 1 (Optional)

Format (18I4) Used if KNDEX1 = 01 (Cols. 26–27 on the Problem Card)
The same instructions apply for this option as above, except that these indices apply to the new variates and/or covariates

continued

after linear transformation No. 1. NTVAR1 + NTCOV1 (Cols. 42–43, Cols. 44–45 on the Problem Card) is the number of indices read in, 18 per card.

Special Index Card(s) for New Variates and Covariates after Linear Transformation No. 2 (Optional)

Format (18I4) Used if KNDEX2 = 01 (Cols. 28–29 on the Problem Card)
The same instructions apply for this option as above, except that these indices apply to variables after linear transformation No. 2. NTVAR2 + NTCOV2 (Cols. 48–49, Cols. 50–51 on the Problem Card) is the number of indices read in, 18 per card.

Algebraic Transformations Card (Optional)

Format (18I4) Used if ITRANS (Cols. 30–31 on the Problem Card) is non-blank.
If algebraic transformations of the data are desired they can be indicated in the following manner.

Columns	*Contents*
1– 4	Index of variable to be transformed
5– 8	Code for transformation of same
9–12	Index of variable to be transformed
13–16	Code for transformation of same
17–20 . : :	Etc. for 72 columns of each card. ITRANS (Cols. 30–31 on the Problem Card) is the number of transformations. If more than 9 are needed, continue punching them 9 per card.

It is important to realize that the algebraic transformations are performed *after* linear transformation No. 1. Thus the index numbers of the variables which are to be algebraically transformed must conform to the index numbers of the "new" variates and covariates after linear transformation No. 1. However, if linear transformation No. 1 is not performed then the index numbers of the original variates and covariates should be used.

It is not required that the transformations all be done on different variables. If desired, the same variable may be transformed as many times as desired.

Algebraic Data Transformations

Ten fixed options and one flexible option for transformations of data are available.

continued

646

Code	*Transformation*	*Restriction*
01	$y = 1/x$	$x \neq 0$
02	$y = \sqrt{x}$	$x \geq 0$
03	$y = \log_e x$	$x > 0$
04	$y = \sin x$	$\|x\| < 2.2 \times 10^{14}$, x in radians
05	$y = \cos x$	$\|x\| < 2.2 \times 10^{14}$, x in radians
06	$y = \tanh x$	
07	$y = e^x$	$x \leq 740$
08	$y = \arcsin x$	$\|x\| \leq 1$
09	$y = \arctan x$	
10	$y = \text{arctanh } x$	$\|x\| < 1$
11	$y = x + C$	

Note: The constant C must be supplied for each use of algebraic transformation No. 11.

These transformations may be applied to separate variables or successively to the same variable; i.e.,

$$y = \sqrt{\frac{1}{x_1}}\ , \quad y_2 = x_2 \quad \text{or} \quad y_1 = \frac{1}{x_1}\ .$$

Constant Card(s) for Algebraic Transformation No. 11 (Optional)

Format to be supplied. (One Format Card)
If algebraic transformation 11 is used, the constant must be supplied for each use of this transformation. The constants must appear on the card or cards in the same order as they are to be used. In addition a Format Card must be supplied for the constants on the card(s). The Format Card should come immediately before the Constant Card or cards for transformation No. 11. If there are N uses of transformation No. 11, N values are immediately read using the supplied format.

Augmenting the Within-cells Sum of Squares to Obtain the Error Sum of Squares (Optional)

Format (18I4) Used if MORE = 01 (Cols. 32–33 on the Problem Card). If it is desired to augment the within-cells sum of squares from the between sums of squares this may be done. This allows the program to be used for Latin squares and for pooling of interaction terms with error. First, the columns of the design

continued

matrix must be identified. (If necessary, the columns of this matrix must be reordered using the ordering feature so that the sources to be pooled appear adjacent to one another). Then the card must be made out as follows.

Columns	*Contents*
1– 4	The column number of the first contrast to be used to augment error.
5– 8	The column number of the last of the contrast to be pooled with error.

Regrouping Statistical Tests Card(s) (Optional)

Format (18I4) Used if IGROUP = 01 (Cols. 34–35 on the Problem Card).
The natural grouping of degrees of freedom or contrasts occurs when the natural ordering of statistical tests is to be used. An example of this "convention" is listed below for a three factor design.

Conventional Order	Effect	D.F.
1	Grand Mean	1
2	C	df(C)
3	B	df(B)
4	BXC	df(B)Xdf(C)
5	A	df(A)
6	AXC	df(A)Xdf(C)
7	AXB	df(A)Xdf(B)
8	AXBXC	df(A)Xdf(B)Xdf(C)

It will be noted that in an N factor design there are "conventionally" 2^N effects.

In some designs, which are essentially factorial, it sometimes is desirable to "pool" several interaction terms and test them simultaneously or to "partition" the degrees of freedom for some effect into two or more parts. As an example, a study of drugs was made involving a factor with four nominal classes: one placebo and three active drugs. It makes good sense to partition this effect into two groups: (a) placebo contrasted with the three drugs and (b) the three drugs contrasted with each other. Aside from the problem of writing appropriate contrasts for the effects, the three degrees of freedom of this factor can be partitioned into *one degree of freedom* for the placebo versus the drugs and *two degrees of freedom between drugs* themselves. Thus we would group the design contrasts into two groups of 1 and 2 d.f. rather than one group of 3 d.f.

Columns	*Contents*
1– 4 5– 8 ⋮ 69–72	Enter the degrees of freedom (right-justified) in the order in which the contrasts appear in the design matrix. NTEST (Cols. 21–23 on the Problem Card) is the number of fields to be punched, 18 per card.

The reader may find some advantage in understanding this option by reading the following section.

Special Orderings of Statistical Tests (Optional)

Format (18I4) Used if KORDER = 01 (Cols. 36–37 on the Problem Card).
The "conventional ordering" of the orthogonal effects explained in the section on statistical tests cards above is GM, C, B, BXC, A, AXC, AXB, AXBXC.

With this feature, it is possible to rearrange this order to any desired order such as GM, A, B, AXB, C, AXC, BXC, AXBXC. Furthermore, it is possible to break each cluster of orthogonal estimates, say B, down into smaller clusters and arrange them into some unusual order. Suppose A has one degree of freedom, B has 3 and C has 2. Associated with these degrees of freedom are contrasts C_1, C_2, B_1, B_2, B_3, A_1. The interaction terms also have contrasts and which we shall label B_1XC_1, B_1XC_2, B_2XC_1, B_2XC_2, etc. It is quite possible that an experimenter would like to cluster these effects in some unusual manner by using the Statistical Tests Cards and also to order them in an unusual manner. For this purpose it is best to locate the column number of the orthogonal contrast in the "conventional order" as is done in the first two columns of the following table. Then, determine in what order these columns should finally appear as done in Column 3.

In this table it is supposed that contrasts B_2 and B_3 are in some way separable from contrast B_1 and that separate tests on B_2 and B_3 are desired. In addition, all tests involving C are run first (they appear at the end of the design matrix).

It is important to realize in this connection that when one has a non-orthogonal design that the ordering of the tests is of

continued

great importance if a least squares analysis is to be run, i.e., if any of the factors are to be weighted. The tests of greatest interest should be run first (appear at the end of the design matrix).

Conventional Order Column	Orthogonal Estimate	New Column
1	grand mean	1
2	C_1	9
3	C_2	10
4	B_1	3
5	B_2	4
6	B_3	5
7	B_1C_1	11
8	B_1C_2	12
9	B_2C_1	13
10	B_2C_2	14
11	B_3C_1	15
12	B_3C_2	16
13	A_1	2
14	A_1C_1	17
15	A_1C_2	18
16	A_1B_1	6
17	A_1B_2	7
18	A_1B_3	8
19	$A_1B_1C_1$	19
20	$A_1B_1C_2$	20
21	$A_1B_2C_1$	21
22	$A_1B_2C_2$	22
23	$A_1B_3C_1$	23
24	$A_1B_3C_2$	24

This ordering will allow the statistics to be computed for the following tests (let D represent the contrasts B_2 and B_3): G.M., A, B_1, D, C, B_1XC, DXC, AXC, AXBXC, AXDXC.

To accomplish this ordering, merely list column 3 of the table on the Special Orderings Cards.

Columns	*Contents*
1– 4	0001
5– 8	0009
9–12	0010
13–16	0003
17–20	0004

continued

Columns	*Contents*
21–24	0005
25–28	0002
29–32	0011
⋮	Etc. for 72 columns of the card

Continue on as many cards as necessary punching 18 per card. NCELLS (Cols. 18–20 on the Problem Card) is the number of items to be punched. *Note:* It is necessary to use the Statistical Tests Cards to make the degrees of freedom conform to the special ordering as it will finally appear.

Special One-Way Contrasts and Special Design Matrices (Optional)

Format must be supplied.
It may often happen that an experimenter either does not wish to analyze his data using the one-way contrast matrices built into the program or does not wish to use the Kronecker product scheme to develop the design basis. For this reason this option is provided.

a. Special one-way Contrast Matrices on *cell means* (Used if the second digit in a contrast option code is 3.)

After the one-way contrast matrix is decided upon, it should be written out as an (m X c) matrix where m = c = the number of levels of the factor in question (i.e. the level from the Contrast Option Card). A single Format Card must be supplied before these cards. For a three level factor with a sample contrast, the (m X c) matrix is shown below (rows are the cell means and columns are the contrasts).

$$\begin{array}{c|ccc} & C_1 & C_2 & C_3 \\ \hline \bar{x}_1 & 1 & 2 & 0 \\ \bar{x}_2 & 1 & -1 & 1 \\ \bar{x}_3 & 1 & -1 & -1 \end{array}$$

The input cards would be punched as shown:

```
          Col 1234567
     Card
       1       (3F2.0)
       2        1 2 0
       3        1-1 1
       4        1-1-1
```

continued

b. Special contrast matrices on *parameters*. (Used if the second digit in a contrast option code is 4.)

After the one-way contrast matrix on the parameters is decided upon it should be written out as a (c X p) matrix where c designates contrast and p designates parameters. *Note:* the number of contrasts will be equal to the number of levels for the factor and the number or parameters would be equal to the number of levels of the factor plus one (i.e. the level from the Contrast Option Card). For three-level factor a (c X p) matrix is shown below. A single Format Card must be supplied before these cards. The input cards would be punched as shown below.

$$\begin{array}{c|cccc} & \mu & \alpha_1 & \alpha_2 & \alpha_3 \\ \hline c_1 & 1 & 0 & 0 & 0 \\ c_2 & 0 & 1 & 1 & -2 \\ c_3 & 0 & 1 & -2 & 1 \end{array}$$

```
      Col. 12345678
Card
  1        (4F2.0)
  2         1 0 0 0
  3         0 1 1-2
  4         0 1-2 1
```

c. Special bases for design matrices (Used if the second digit in a contrast option code is 5.)

When non-factorial designs or other designs where Kronecker products of one-way design matrices are not appropriate are used, it is possible to fabricate an appropriate design matrix and insert this into the program.

The special design basis matrix should be laid out as an (m X c) matrix having cell means appearing as rows and contrasts as columns. The matrix is then punched on cards the same way that special one-way contrast matrix on cell means is (see a. above), *preceded by a single appropriate Format Card.*

continued

Interval Width Card(s) for Contrast Option No. 07 (Optional)

An appropriate format must be supplied for the card(s) on one Format Card. (Used if the second digit in a contrast option code is a 7.)

If contrast option No. 07 is used it is necessary to specify the relation "distance" between the treatment levels of the given factor. For example a treatment factor might consist of different amounts of practice time, i.e. 5, 10, 15, 30 minutes.

One would enter on a card the numbers: 1, 2, 3, 6 which conveniently express the relative size of the treatments.

Linear Transformation No. 1 Cards (Optional)

A single Format Card must be supplied for these cards. (Used if LT1 = 01, Cols. 40–41, on the Problem Card.)
If linear transformation No. 1 is desired, a Pn X Po matrix of coefficients must be supplied at this point; where Po equals the number of original variates and covariates (IVAR + ICOVAR, Cols. 10–11, Cols. 12–13 on the Problem Card) and Pn equals the number of new variates and covariates (NTVAR1 + NTCOV1, Cols. 42–43, Cols. 44–45 on the Problem Card). This feature can be used for reordering the variates and covariates in order to satisfy the requirement that all variates precede covariates. Each linear transformation of the original variates corresponding to a new variate of covariate must begin on a new card, i.e. each new card is a new variate.

Cell Frequency Cards (N–Cards) and Data Cards

Format for the N–Cards is (12X,I6); format for the Data Cards was supplied above, on card following Test Titles Cards.

In order to arrange the data for input, the data must be arranged by cells of the design and entered in the natural numerical order. For example, let us consider the following experiment: A having 2 levels, B having 3 levels and C having 2 levels. The natural numerical order of the cells of the design is shown below. The data for each cell *must be preceded by an N–Card* denoting how many observations are included in that cell and *must have* this number in Cols. 13–18 punched to the far right of the field. Cols. 1–12 are available to the user. The layout of this card is not flexible.

continued

Level of A	Level of B	Level of C
1	1	1
1	1	2
1	2	1
1	2	2
1	3	1
1	3	2
2	1	1
2	1	2
2	2	1
2	2	2
2	3	1
2	3	2
:	:	:
	etc.	

If a cell has no observations in it and is anticipated in the bases of the design, an N–Card must be supplied.

It will be noted that this program does *not* allow for missing observations for some variables. Each observation *must have* a complete set of scores on all the variates and covariates.

Linear Transformation No. 2 Cards (Optional)

A single Format Card must be supplied. (Used if LT2 = 01, Cols. 46–47 on the Problem Card.)
The same identical instructions hold for this option as for linear transformation No. 1. Again note that a single Format Card must be supplied before the Linear Transformation Cards. In this case, Po = IVAR + ICOVAR, Cols. 10–11, Cols. 12–13 of the Problem Card and Pn = NTVAR2 + NTCOV2, Cols. 48–49, Cols. 50–51 on the Problem Card.

However, the use to which it is put will probably be different, i.e., if one has p variates and wishes to test whether they are affected *differently* by a set of treatments, one could set up P–1 orthogonal linear combinations of the original variates and proceed with the analysis. Each effect tested could be interpreted as an "Interaction" between variates and the particular treatment. Variates can be thought of as a "factor".

Recombination Analysis Cards (Optional)

Format (18I4) (Used if NSKP = 01, Cols. 54–55 on the Problem Card.)

continued

This feature of the program is designed to allow, in the same computer run, an analysis of either a subset of the original variates and covariates and/or a permutation of variates and covariates.

Use of this feature requires at least two cards. The first card contains the number of variates and covariates involved in the new analysis. The number of variates is punched in Cols. 1–4 to the extreme right of the field; the number of covariates is punched in Cols. 5–8 to the extreme right of the field.

The second card (or remaining cards) is to be punched with the index numbers of these variables, in four-column fields. The number of items is the number of variates + number of covariates from the previous card. *Note:* Proper indices *must* be used.

There is no limit to the number of recombinations which may be made.

End of Recombination Card (Required)

Format (9999999999)
When all the Recombination Cards have been entered into the deck, or if there are no Recombination Cards, a card with nines in Cols. 1–10 must follow to signify that no other recombinations are to be expected.

Last Problem Card

This program will handle any number of separate problems with attendant *recombination analyses*. To signify that no more problems appear, a card with 9's in Cols. 1–10 must be placed after the last problem's last End of Recombination Card (or instead, a Finish Card, FINISH in Cols. 1–6 may be used). This means that two (2) "nines" cards or one "nines" card followed by a Finish Card end each UMST570 deck, (one End of Recombination Card and one Last Problem Card or one Finish Card).

Order of Cards

For each problem

Problem Card (Required)
Problem Title Card(s) (Optional)
Contrast Option Card(s) (Required)
Test Title Card(s) (Required)

continued

Variable Format Card (Required)
Special Index Cards for Original Data (Optional)
Special Index Cards for New Data after Linear Transformation No. 1 (Optional)
Special Index Cards for New Data after Linear Transformation No. 2 (Optional)
Algebraic Data Transformations Card (Optional)
Constant Card(s) for Algebraic Transformation No. 11 (Optional)
Augmenting Within-Cells Sum of Squares for Error (Optional)
Regrouping Statistical Tests Card(s) (Optional)
Special Ordering of Statistical Tests (Optional)
Special One-Way Contrasts and Design Matrices (Optional)
Interval Width Card(s) for Contrast Option No. 07 (Optional)
Linear Transformation No. 1 Card(s) (Optional)
N-Cards and Data Cards (Required)
Linear Transformation No. 2 Card(s) (Optional)
Recombination Analysis Cards (Optional)
End of Recombination Card (Required)

After last problem

Finish Card or Last-Problem Card (Required)

SAMPLE INPUT

```
PROBLEM 1 3   2 3  6  4                                 1   2
THREE UNIVARIATE ANALYSES OF VARIANCE ON A 2-WAY CROSSED CLASSIFICATION,FACTOR
A HAVING 3-LEVELS, AND FACTOR B HAVING 2-LEVELS
 3 2 2 2
   GM
   B
   A
   AB
(3F3.0)
                 3
  2  3  4
  5  3  8
  2  1  3
                 4
  7  6  1
  5  2  3
  1  6  9
  3  9  5
                 7
  1  5  6
  9  7  2
  2  0  6
  2  8  2
  1  7  8
  8  5  1
  7  3  6
                 3
  3  1  3
  3  7  2
  6  9  6
                 3
  2  6  0
  5  4  1
  5  6  9
                 7
  6  1  4
  8  9  1
  5  4  4
  2  1  1
  2  1  1
  6  9  4
  8  1  1
   1
   1
   1
   2
   1
   3
9999999999
9999999999
```

SAMPLE OUTPUT

```
UMST570 -- U OF M VERSION OF MANOVA

MULTIVARIATE ANALYSIS OF VARIANCE, GENERAL LINEAR HYPOTHESIS MODEL

PROBLEM NUMBER   1

THIS PROBLEM REQUIRES A FIELD LENGTH OF 050400B

THREE UNIVARIATE ANALYSES OF VARIANCE ON A 2-WAY CROSSED CLASSIFICATION,FACTOR

A HAVING 3-LEVELS, AND FACTOR B HAVING 2-LEVELS

CONTROL CARDS
 1 3-0 2 3  6  4-0-0-0-0-0-0-0-0-0-0-0-0-0 1-0 2-0-0-0*
 3 2 2 2
   GM
   B
   A
   AB
(3F3.0)


THERE ARE          3    OBSERVATIONS IN CELL    1
THERE ARE          4    OBSERVATIONS IN CELL    2
THERE ARE          7    OBSERVATIONS IN CELL    3
THERE ARE          3    OBSERVATIONS IN CELL    4
THERE ARE          3    OBSERVATIONS IN CELL    5
THERE ARE          7    OBSERVATIONS IN CELL    6


THE MEANS FOR EACH CELL, LISTED AS ENTERED
VAR.       1               2              3
CELL I D
 1 1
                    WITH    3 OBSERVATIONS
        3.0000         2.3333         5.0000
 1 2
                    WITH    4 OBSERVATIONS
        4.0000         5.7500         4.5000
 2 1
                    WITH    7 OBSERVATIONS
        4.2857         5.0000         4.4286
 2 2
                    WITH    3 OBSERVATIONS
        4.0000         5.6667         3.6667
 3 1
                    WITH    3 OBSERVATIONS
        4.0000         5.3333         3.3333
 3 2
                    WITH    7 OBSERVATIONS
        5.2857         3.7143         2.2857


MARGINAL MEANS, ZERO DESIGNATES SUBSCRIPT SUMMED OVER
VAR.       1               2              3
MARG. I D
 0 1
                    WITH     13.000 OBSERVATIONS
        3.9231         4.4615         4.3077
 0 2
                    WITH     14.000 OBSERVATIONS
        4.6429         4.7143         3.2143
 1 0
                    WITH      7.000 OBSERVATIONS
        3.5714         4.2857         4.7143
 2 0
                    WITH     10.000 OBSERVATIONS
        4.2000         5.2000         4.2000
 3 0
                    WITH     10.000 OBSERVATIONS
        4.9000         4.2000         2.6000
 0 0
                    WITH     27.000 OBSERVATIONS
        4.2963         4.5926         3.7407
```

continued

```
BASIS OF DESIGN WITH EFFECTS EQUATIONS IN COLUMNS
EFFECTS           1          2          3          4          5          6
CELL I D
 1 1
               1.000      1.000      1.000      0.         1.000      0.
 1 2
               1.000     -1.000      1.000      0.        -1.000      0.
 2 1
               1.000      1.000     -.5000      1.000     -.5000      1.000
 2 2
               1.000     -1.000     -.5000      1.000      .5000     -1.000
 3 1
               1.000      1.000     -.5000     -1.000     -.5000     -1.000
 3 2
               1.000     -1.000     -.5000     -1.000      .5000      1.000


THIS PROBLEM REQUIRES A FIELD LENGTH OF 050300B
THE ERROR CORRELATION MATRIX

 VAR.     1        2        3
   1
      1.0000    .1835  -.1989
   2
       .1835  1.0000    .0966
   3
      -.1989    .0966  1.0000


THE PRINCIPAL COMPONENTS (COLUMNWISE) OF THE ERROR CORRELATION MATRIX
COMP.    1 ST    2 ND    3 RD

VAR.
   1
        .8482   .0402   .5282
   2
        .4512   .7786  -.4362
   3
       -.5514   .6989   .4556

ROOTS
       1.2269  1.0962   .6769
```

continued

```
U OF M VERSION OF MANOVA
MULTIVARIATE ANALYSIS OF VARIANCE, GENERAL LINEAR HYPOTHESIS MODEL
UNIVARIATE ANALYSIS OF VARIATE      1, WITH -0 COVARIATES
                              DF    SUM OF SQUARES    MEAN SQUARES   F-RATIO           P-VALUE
ANOVA ERROR                   21      150.85714       7.1836735
AB                             2      2.8972050       1.4486025      .20165205       .81894354
A                              2      6.3830148       3.1915074      .44427234       .64718029
B                              1      3.4922670       3.4922670      .48613944       .49330191
GM                             1      498.37037       498.37037      69.375421       .00000004
```

```
U OF M VERSION OF MANOVA
MULTIVARIATE ANALYSIS OF VARIANCE, GENERAL LINEAR HYPOTHESIS MODEL
UNIVARIATE ANALYSIS OF VARIATE      2, WITH -0 COVARIATES
                              DF    SUM OF SQUARES    MEAN SQUARES   F-RATIO           P-VALUE
ANOVA ERROR                   21      196.17857       9.3418367
AB                             2      23.934472       11.967236      1.2810367       .29858005
A                              2      7.9748686       3.9874343      .42683622       .65810636
B                              1      .43060643       .43060643     4.60944076E-02 .83207546
GM                             1      569.48148       569.48148      60.960333       .00000012
```

```
U OF M VERSION OF MANOVA
MULTIVARIATE ANALYSIS OF VARIANCE, GENERAL LINEAR HYPOTHESIS MODEL
UNIVARIATE ANALYSIS OF VARIATE      3, WITH -0 COVARIATES
                              DF    SUM OF SQUARES    MEAN SQUARES   F-RATIO           P-VALUE
ANOVA ERROR                   21      165.47619       7.8798186
AB                             2      .28516218       .14258109     1.80944636E-02 .98208355
A                              2      17.365021       8.6825105      1.1018668       .35070742
B                              1      8.0588116       8.0588116      1.0227154       .32338681
GM                             1      377.81481       377.81481      47.947146       .00000077
```

COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 0.332 seconds. At the current rate for the Univ. of Minnesota ($0.20/sec.), the computer time cost $0.07 plus a charge for output supplies.

Charge to user = computer costs + postage + network overhead
= $0.29 + postage + network overhead

CONTENTS—UMST570

pages	
1– 2	Identification & Abstract
3–21	User Instructions
23–26	I/O
27	Cost—Contents

661

000 0144

DESCRIPTIVE TITLE	Stepwise Regression
CALLING NAME	UMST580
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Mr. M. Dale Fimple Sandia Corporation
LANGUAGE	CDC Fortran IV Modified for use at Univ. of Minnesota
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program performs stepwise linear regression handling as many as 80 independent and 25 dependent variables. There is an option for using variable weights on the observations. Thirty-four methods of transformation and/or generation of individual variables as possible. This program selects a subset of independent variables by fixed F or fixed probability, deleting least significant variables one at a time.

Output includes regression and correlation parameters at each step. An optional feature of this program can compute and give as output residual information. Since the data analyzed may differ from the input data, all or part of the data analyzed may also be included as output.

Any or all of the following may be calculated and given as output:

1. Sums, sums of squares, and sums of cross products
2. Means, sums of squares, and sums of cross products
3. Correlation coefficients
4. The inverse of the correlation matrix and the inverse of the X'X matrix.

continued

662

Method

The method used by UMST580 to pick the most significant subset of independent variables for each dependent variable is as follows:

1. The regression analysis is performed on the entire set of P independent variables.
2. For each variable in the analysis, the statistic, b_i^2/C_{ii}, is computed where b_i is the regression coefficient giving the relationship between the ith independent variable and the dependent variable. C_{ii} is the ith diagonal element of $(X'X)^{-1}$. This statistic gives the reduction in the regression sum of squares when X_i is deleted from the analysis.
3. The minimum of b_i^2/C_{ii} is obtained.
4. This value is divided by the error mean square at that point, and this ratio has an F distribution.
5. This value of F is tested against either a value of F or a probability level included in the control card for the problem.
6. If it is determined that this minimum sum of squares is significant, the deletion process is discontinued. Otherwise, X_i is deleted, the inverse matrix is adjusted for the deletion, and a branch is taken back to step 2.

Since the method given above for selecting the optimum subset of variables is equivalent to partially reinverting the X'X matrix, it is essential to include a check on the accuracy of the inverse. It should be pointed out that all operations are carried out on the correlation matrix which has all elements between ± 1. The basic inversion routine used throws out any variables which would cause singularity. After the correlation matrix is inverted, the norm of $(I - RB_0)$ is computed. I is the P X P identity matrix, R is the correlation matrix, and B_0 is the computed estimate of R^{-1}. The norm used is defined by

$$N(A) = \sqrt{\sum_i \sum_j a_{ij}^2}$$

If this norm does not meet the requirements specified on the control card, the option is available to use Hotelling's method[1] to obtain a better estimate of the inverse. This is an iterative technique, and the ith estimate of the inverse is given by

$$B_i = B_{i-1}(2I - RB_{i-1}).$$

continued

If the norm mentioned above is less than 1.0, convergence to the true inverse is assured (in theory). Actually, the degree of convergence is restricted by the use of floating point arithmetic.

The following table[2] gives probability values computed by UMST580 for selected F ratios with one degree of freedom in the numerator and the number indicated in the denominator. The true probability values are given across the top.

Degress of Freedom	True Probability			
	0.70	0.90	0.95	0.99
1	0.6927	0.8687	0.9073	0.9386
2	0.6988	0.8950	0.9422	0.9806
3	0.6991	0.8992	0.9477	0.9869
4	0.7001	0.9005	0.9497	0.9887
5	0.7000	0.9014	0.9505	0.9895
6	0.7000	0.9020	0.9510	0.9897
7	0.6989	0 9022	0.9512	0.9899
8	0.6997	0.9025	0.9515	0.9902
9	0.6994	0.9026	0.9517	0.9903
10	0.7003	0.9025	0.9517	0.9901
12	0.6986	0.9031	0.9520	0.9903
14	0.6996	0.9030	0.9520	0.9903
16	0.6998	0.9033	0.9520	0.9903
18	0.6996	0.9035	0.9521	0.9903
20	0.6989	0.9032	0.9522	0.9903
40	0.6965	0.9041	0.9524	0.9903
100	0.6981	0.9043	0.9528	0.9903
500	0.7002	0.9045	0.9527	0.9903

REFERENCES

1. Hotelling, H., "Some New Methods in Matrix Calculation," Annals of Math. Stat., *14* (1943), pp. 1—34.

2. Kendall, M.G., *The Advanced Theory of Statistics* (2nd ed.; London: Griffin Pub. Co., 1967), *II*, Equation 21.62, p. 118.

USER INSTRUCTIONS

System Control Cards

System Control Cards will be prepared by Univ. of Minnesota personnel.

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM

Columns 8–11 and 12–16 are mutually exclusive. That is, you may choose to delete insignificant variables from the regression on a constant probability basis *or* on a constant F-ratio basis.

8–11	ALFA	Probability level expressed as the area under the F curve from zero to F_0, the significant value of F. The decimal point is assumed to be to the left of Col. 8 unless it is punched. 0: otherwise
12–16	FSIG	F ratio with the decimal point assumed between Cols. 14–15 unless it is punched. 0: otherwise
17–21	NOB	Number of observations
22–26	P	Number of independent variables to be included in analysis, $P \leq 80$
27–31	Q	Number of dependent variables to be included in analysis, $Q \leq 25$
32–36	N	Number of independent variables punched on Input Data Card(s), $N \leq 800$
37–41	M	Number of dependent variables punched on Input Data Card(s) $M \leq 100$
42–44	NITMAX	006: if errors in the inverse of the correlation matrix are greater than those specified, an iterative technique is applied to improve the inverse. 000: no iteration

continued

Columns	*Parameter*	*Contents*
45–46		06
47–48	IFHED	01: heading information desired (see Heading Cards) 00: otherwise
49–50	NCON	Number of constants, NCON $\leq$ 97 (see Constant Cards)
51–52	IFWT	01: weighted regression analysis 00: otherwise
53–54	INWT	(To be used only if Cols. 51–52 = 01) 01: weights included in input data 00: weights to be computed or derived from input data during data transformation
55–56		blank
57–60	TOL	Value, usually 0.001, (decimal point assumed to left of field if not punched), used in testing for singular matrix.
61–62	IFXX	00: output of sums, sums of squares, sums of cross products 01: otherwise
63–64	IFSX	00: output of means, sums of squares, sums of cross products 01: otherwise
65–66	IFCOR	00: output of correlation coefficients 01: otherwise
67–68	IFINV	00: output of inverse of correlation matrix and inverse of X'X matrix 01: otherwise
69–70	IFRES	01: output residual information 00: otherwise
71–72	IFDATO	–1: output all data analyzed 00: otherwise

Heading Card(s)

If Cols. 47–48 of the Problem Card = 01, any number of 80-column Heading Cards will be written without alteration on the output. (Col. 1 is used only for control information to the printer; i.e., 1 = new page, 0 = double space, etc.) The last such card must contain ENDHED in the first six columns.

continued

Format Card(s)

As regards the data, the format statement must provide for reading an ID field, N independent variables ($N < 800$) and M dependent variables ($M \le 100$) (Cols. 32–36, Cols. 37–41 on Problem Card). In this section N and M include variables used in transforming and generating the variables actually used in the regression. (Recall that P and Q, Cols. 22–26, Cols. 27–31 on Problem Card, are restricted to 80 and 25 respectively for the actual regression.) F, E, and X fields are used to read in all variables including the ID field. The ID field is not included as a dependent or independent variable on input.

If the Problem Card has indicated that weights are to be read in with the data, the weight for each observation will be assumed to be in the field following the last dependent variable.

A separate format statement is required for reading in constants.

The layout of a Format Card for constants or data is as follows:

Columns	*Content*
1– 3	The total number of columns including open and close parentheses used in the format statement divided by 6 (plus 1 if there is a remainder).
4– 6	blank
7–72	Format statement beginning with open parenthesis and ending with close parenthesis

If the format statement cannot be contained on one card, it may be continued on subsequent cards starting in Col. 1 of these cards. These columns continued on subsequent cards are included in the total count used to determine value in Cols. 1–3.

Constant Card(s)

If it is desired to use certain constants in any transformations of the data, they may be punched on a series of cards separate from the Data Cards. They must be preceded by Format Card(s) as described above if NCON, (Cols. 49–50 on the Problem Card) is non-blank. The number of constants is NCON.

OP Card(s)

The OP Cards allow for 34 transformations and generations of the input data. Each OP Card contains one instruction in Cols. 1–15.

continued

Columns 16–72 may be used for comments. There *must* be at least one OP Card with operation order OUT to end the set of transgenerations. See below. The layout is as follows:

Columns	*Contents*
1– 3	OP - the operation code
4– 6	A
7– 9	B
10–12	C
13–15	D

The last OP Card must contain the operation code OUT in Cols. 1–3, followed by blanks in Cols. 4–15.

In order to make use of the transgenerations the user must at least understand where his output information is initially stored. This is done as follows:

Location	*Contents*
1–N	independent variables
801-800+M	dependent variables
901-997	any constants from constant cards
998	ID number
999	weight

After the OP instructions have been carried out, the number of independent variables in the regression is determined by Cols. 22–26 of the Problem Card (P), and they *must* be stored consecutively in locations 1 to P; the number of dependent variables is determined by Col. 27–31 (Q) and *must* be stored consecutively in locations 801 to 800+Q. Transformations are done sequentially; thus two variables may be added in one transformation and the square root of the sum may be done in the next transformation.

In the following operation code list the letter A, B, C, D refers to numbers used by the instruction itself while (A), (B), (C), (D) refers to the data in the location referred to by A, B, C, D. (i.e., (A) + B means that the variable in location A is added to the number B).

continued

Code	*Transformation*	
ADD	(C) = (A) + (B)	
SUB	(C) = (A) – (B)	
MPY	(C) = (A) X (B)	
DVD	(C) = (A) / (B)	
ABS	$(B) = \lvert (A) \rvert$	
SQT	$(B) = \sqrt{(A)}$	$(A) \geq 0$
EXP	$(B) = e^{(A)}$	$(A) < 740$
LOG	(B) log(A)	$(A) > 0$
MOV	A items beginning at (B) are moved to the locations beginning at (C). (C) must not be in the range (B) to (B) + A–1	

DUM	(A) is used to generate B dummy variables in locations C through D, [D = C + number of levels of the variable in the location A minus 1], as follows
	if (A) = 0, all dummy variables = 0;
	if (A) = r and $C+r-1 \leq D$, (C+r–1) =1
	and all the rest of the dummy variables = 0	

For codes ADA through LGA the array may be truncated on the right by simply not allowing enough locations to contain the array. For ADA, SUA, MPA, DVA, ABA, SQA, EXP, LGA, A must be < B.

ADA	the data in (A) through (B) are used to form $\frac{[B - A + 1]!}{2![B - A - 1]!}$ two-variable sums which are stored in locations C through D in the form: C is (A) + (A+1), C+1 is (A) + (A+2), etc.
SUA	same as ADA except differences are generated
MPA	same as ADA except products are generated
DVA	same as ADA except quotients are generated (denominator must not be zero).
ABA	absolute values of (A) through (B) are stored in locations C through D
SQA	same as ABA except square roots are obtained. (Values must be ≥ 0)

continued

Code	*Transformation*
EXA	same as ABA except exponentials are obtained. (e^x; values must be < 740)
LGA	same as ABA except natural logarithms are obtained (values must be > 0)
GAS	(A) is assumed to be a normal deviate with mean zero and variance one. The area under the normal curve from $-\infty$ to (A) is obtained and put in location B
STP	(A) is used to generate B dummy "step" variables in locations C through D, as follows: the first A new variables = 1, the remainder = –1
NED	for $0 \leq (A) \leq 1.0$, (A) is assumed to be an area under a normal curve. The normal equivalent deviate corresponding to (A) is put in location B.
PBT	same as NED except the normal equivalent deviate is increased by 5.0 yielding a "PROBIT"
INT	$(B) =$ largest integer $\leq \lvert (A) \rvert$; (B) will have the same sign as (A)
MOD	$(C) =$ remainder upon division of (A) by (B). $(B) > 0$.
MAX	$(C) =$ the larger of (A) and (B)
MIN	$(C) =$ the smaller of (A) and (B)
SIN	$(B) = \sin(A)$; $\lvert (A) \rvert < 2.2 \times 10^{14}$, (A) in radians
COS	$(B) = \cos(A)$; $\lvert (A) \rvert < 2.2 \times 10^{14}$, (A) in radians
ATN	$(B) = \arctan(A)$
THF	$(B) = \tanh(A)$
ASN	$(B) = \arcsin(A)$; $\lvert (A) \rvert \leq 1$
GEX	$(C) = (A)^{(B)}$
DCT	(A) is rounded to an integer, I. If $I = 0$, $(D) = (B)$. If $I \geq 1$, $(D) = (C)$. If $I < 0$, error termination of problem.
POL	It is assumed that $1 \leq (A) \leq B$; B may be thought of as the number of levels of some variable in an analysis of variance, and (A) is the particular level assigned to the observation under question. C is the number of orthogonal polynomials which are stored starting at (D). $C < B-1$.

continued

Code	*Transformation*
OUT	End of set instructions

Finish Card

Columns	*Contents*
1- 6	FINISH

Used after last problem.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

- Program Control Cards for problem 1
  - Problem Card
  - Heading Card(s), if any
  - Format Card(s) for constants, if any
  - Constants Card(s), if any
  - Format Card(s) for data
  - Op Card(s) (at least one card with Op code OUT)
- Data Cards for problem 1
- Program Control Cards for problem 2
- Data Cards for problem 2

:

- Program Control Cards for last problem
- Data Cards for last problem
- Finish Card

SAMPLE INPUT

```
PROBLEM4000          10     5     1     6     0  6 6 1          .001          1-1
   SAMPLE PROBLEM
ENDHED
  5    (F2.0,F5.2,F6.0,2F6.2,2F6.0)
MOV  1  6801
OUT
 1 0250      25  2500    150     34     64
 2 1300      21  2100    087     36     65
 3 0350      22  2200    043     41     82
 4 0175      09  0130    180     15     23
 5 0300      23  2300    200     33     64
 6 0200      10  0060    330     13     16
 7 0550      07  0140    340     16     12
 8 0600      06  0080    500     11     27
 9 0130      08  0270    150     19     48
10 0500      18  0360    180     27     50
FINISH
```

671

SAMPLE OUTPUT

```
UMST580    STEPWISE REGRESSION

PROBLEM CARD IS PRINTED BELOW

              1111111111222222222233333333334444444444555555555566666666667777777778
   1234567890123456789012345678901234567890123456789012345678901234567890123456789O
    PROBLEM4000         10    5    1    6    0  6 6 1          .001          1-1       U5RT1
 SAMPLE PROBLEM                                                                    U5BT1      3


        NUMBER OF OBSERVATIONS =      10
        NUMBER OF INDEPENDENT VARIABLES IN ANALYSIS =       5
        NUMBER OF DEPENDENT VARIABLES IN ANALYSIS   =       1
        NUMBER OF INDEPENDENT VARIABLES IN INPUT DATA =       6
        NUMBER OF DEPENDENT VARIABLES IN INPUT DATA    =       0
        TOLERANCE FACTOR FOR MATRIX INVERSION =      .0010000
        PROBABILITY LEVEL FOR DELETING VARIABLES =  .400000

        SUMS OF X VALUES
   ( 1)   4.355000E+01      ( 2)   1.490000E+02      ( 3)   1.014000E+02      ( 4)   2.160000E+01
   ( 5)   2.450000E+02      (

        RAW XX MATRIX
( 1, 1)  2.9650250E+02       ( 1, 2)  6.9215000E+02       ( 1, 3)  5.1898500E+02       ( 1, 4)  9.1965000E+01
( 1, 5)  1.1614500E+03       (
( 2, 2)  2.7330000E+03       ( 2, 3)  2.1977000E+03       ( 2, 4)  2.5863000E+02       ( 2, 5)  4.3480000E+03
( 3, 3)  2.1039000E+03       ( 3, 4)  1.3484000E+02       ( 3, 5)  3.4740000E+03       (
( 4, 4)  6.3371800E+01       ( 4, 5)  4.2235000E+02       (
( 5, 5)  7.0830000E+03       (

        XY CROSS PRODUCTS

 Y( 1)
        X( 1)  2.096650E+03          X( 2)  8.138000E+03          X( 3)  6.628500E+03          X( 4)  7.478100E+02
        X( 5)  1.329400E+04          X(

        SUMS OF Y SQUARED
   ( 1)   2.560300E+04      (

        SUMS OF Y VALUES
   ( 1)   4.510000E+02      (

        WEIGHTED DEGREES OF FREEDOM=     10



        MEAN VALUES OF X
   ( 1)   4.355000E+00      ( 2)   1.490000E+01      ( 3)   1.014000E+01      ( 4)   2.160000E+00
   ( 5)   2.450000E+01      (
        SUM (XX).CORRECTED FOR MEANS
( 1, 1)  1.0684225E+02       ( 1, 2)  4.3255000E+01       ( 1, 3)  7.7388000E+01       ( 1, 4) -2.1030000E+00
( 1, 5)  9.4475000E+01       (
( 2, 2)  5.1290000E+02       ( 2, 3)  6.8684000E+02       ( 2, 4) -6.3210000E+01       ( 2, 5)  6.9750000E+02
( 3, 3)  1.0757040E+03       ( 3, 4) -8.4184000E+01       ( 3, 5)  9.8970000E+02       (
( 4, 4)  1.6715800E+01       ( 4, 5) -1.0685000E+02       (
( 5, 5)  1.0805000E+03       (
        SUM (XY).CORRECTED FOR MEANS

 Y( 1)
        X( 1)  1.325450E+02          X( 2)  1.418100E+03          X( 3)  2.055360E+03          X( 4) -2.263500E+02
        X( 5)  2.244500E+03          X(
        SUM (YY).CORRECTED FOR MEANS
   ( 1)   5.262900E+03      (
        MEAN VALUES OF Y
   ( 1)   4.510000E+01      (
```

continued

```
        XY CORRELATION COEFFICIENTS

 Y( 1)
            X( 1)   .176758            X( 2)    .863133             X( 3)    .863831            X( 4)  -.763140
            X( 5)   .941228            X(

        CORRELATION MATRIX
( 1, 1)  1.0000000E+00      ( 1, 2)  1.8477723E-01      ( 1, 3)  2.2827362E-01      ( 1, 4) -4.9762728E-02
( 1, 5)  2.7805647E-01      (
( 2, 2)  1.0000000E+00      ( 2, 3)  9.2468293E-01      ( 2, 4) -6.8266220E-01      ( 2, 5)  9.3694797E-01
( 3, 3)  1.0000000E+00      ( 3, 4) -6.2779750E-01      ( 3, 5)  9.1800446E-01      (
( 4, 4)  1.0000000E+00      ( 4, 5) -7.9505753E-01      (
( 5, 5)  1.0000000E+00      (
        DETERMINANT=   3.2918988*10**    -3

        REQUIRED NORM=   5.000000E-06
        NORM(     1)=   6.241462E-13
        DESIRED INVERSE OBTAINED DURING ITERATION NO.     1
        INVERSE OF CORRELATION MATRIX
( 1, 1)  1.3582070E+00      ( 1, 2)  1.1156399E+00      ( 1, 3)  3.7573487E-01      ( 1, 4) -9.2553338E-01
( 1, 5) -2.5037333E+00      (
( 2, 2)  1.1526399E+01      ( 2, 3) -3.5820155E+00      ( 2, 4) -1.4765574E+00      ( 2, 5) -8.9954886E+00
( 3, 3)  9.2856371E+00      ( 3, 4) -2.1449873E+00      ( 3, 5) -6.9779579E+00      (
( 4, 4)  3.9942703E+00      ( 4, 5)  6.7855906E+00      (
( 5, 5)  2.1925216E+01      (
        INVERSE OF XPRIMEX MATRIX
( 1, 1)  1.2712265E-02      ( 1, 2)  4.7658039E-03      ( 1, 3)  1.1083160E-03      ( 1, 4) -2.1900649E-02
( 1, 5) -7.3689257E-03      (
( 2, 2)  2.2472994E-02      ( 2, 3) -4.8224165E-03      ( 2, 4) -1.5946685E-02      ( 2, 5) -1.2083591E-02
( 3, 3)  8.6321489E-03      ( 3, 4) -1.5996123E-02      ( 3, 5) -6.4724629E-03      (
( 4, 4)  2.3895179E-01      ( 4, 5)  5.0490734E-02      (
( 5, 5)  2.0291731E-02      (


 REGRESSION ANALYSIS ON Y( 1)

           ANOVA

 SOURCE         SUM OF SQUARES     DEGREES OF FREEDOM

 TOTAL          5.26290000E+03         9
 REGRESSION     4.74195556E+03         5
 ERROR          5.20944442E+02         4


 VARIANCE              = 1.30236110E+02
 STANDARD DEVIATION = 1.14121037E+01
 CORRELATION COEF.   = 9.49218471E-01
 R SQUARED             = 9.01015706E-01

             COEFFICIENT      STANDARD ERROR              T RATIO              F RATIO      SUM OF SQUARES
   ( 0)      -6.8388448E+00      3.3008716E+01           -.2071830            .0429248
   ( 1)      -8.6101962E-01      1.2866996E+00           -.6691691            .4477873        5.8318073E+01
   ( 2)      -9.2325407E-01      1.7107879E+00           -.5396660            .2912394        3.7929885E+01
   ( 3)       1.4368595E-01      1.0602912E+00            .1355156            .0183645        2.3917165E+00
   ( 4)       8.4510838E-01      5.5785438E+00            .1514926            .0229500        2.9889217E+00
   ( 5)       2.7005168E+00      1.6256433E+00           1.6611988           2.7595816        3.5939718E+02

 REQUIRED PROBABILITY= .400000
 PROBABILITY FOR X( 3)= .270070
 DELETE X( 3)

           ANOVA

 SOURCE         SUM OF SQUARES     DEGREES OF FREEDOM

 TOTAL          5.26290000E+03         9
 REGRESSION     4.73956384E+03         4
 ERROR          5.23336158E+02         5


 VARIANCE              = 1.04667232E+02
 STANDARD DEVIATION = 1.02307004E+01
 CORRELATION COEF.   = 9.48979061E-01
 R SQUARED             = 9.00561257E-01
```

continued

```
            COEFFICIENT      STANDARD ERROR            T RATIO            F RATIO      SUM OF SQUARES
 ( 0)      -9.7122524E+00     2.2679049E+01           -.4282478           .1833962
 ( 1)      -8.7946804E-01     1.1470237E+00           -.7667392           .5878890       5.1532717E+01
 ( 2)      -8.4298281E-01     1.4388204E+00           -.5858847           .3432608       3.5928161E+01
 ( 4)       1.1113710E+00     4.6805827E+00            .2374429           .0563791       5.9010458E+00
 ( 5)       2.8082538E+00     1.2711875E+00           2.2091578          4.8803783       5.1081569E+02

REQUIRED PROBABILITY= .400000
PROBABILITY FOR X( 4)= .344249
DELETE X( 4)

         ANOVA

SOURCE         SUM OF SQUARES     DEGREES OF FREEDOM

TOTAL          5.26290000E+03          9
REGRESSION     4.73366280E+03          3
ERROR          5.29237204E+02          6


VARIANCE                = 8.82062006E+01
STANDARD DEVIATION = 9.39181562E+00
CORRELATION COEF.   = 9.48388108E-01
R SQUARED               = 8.99440004E-01
            COEFFICIENT      STANDARD ERROR            T RATIO            F RATIO      SUM OF SQUARES
 ( 0)      -4.7312901E+00     7.9120436E+00           -.5979858           .3575871
 ( 1)      -7.7408726E-01     9.7095548E-01           -.7972428           .6355961       5.6063515E+01
 ( 2)      -7.1086133E-01     1.2180710E+00           -.5835959           .3405842       3.0041641E+01
 ( 5)       2.6038479E+00     8.5863077E-01           3.0325584          9.1964102       8.1118040E+02

ALL VARIABLES HAVE PROBABILITIES EQUAL TO OR GREATER THAN  .400000



                  RESIDUAL INFORMATION


RUN NO.      Y( 1)         PREDICTION      DIFFERENCE     DIF/STD DEV.        WEIGHT      STD. ERROR    DIF/STD.ERR
      1   6.400000E+01   6.409279E+01  -9.278764E-02      -.009880         1.000000      11.351700      -.008174
      2   6.500000E+01   6.401601E+01   9.839875E-01       .104771         1.000000      12.650930       .077780
      3   8.200000E+01   8.367822E+01  -1.678220E+00      -.178690         1.000000      12.136578      -.138278
      4   2.300000E+01   2.657402E+01  -3.574024E+00      -.380547         1.000000      10.354927      -.345155
      5   6.400000E+01   6.252362E+01   1.476381E+00       .157199         1.000000      10.690141       .138107
      6   1.600000E+01   2.046195E+01  -4.461945E+00      -.475089         1.000000      10.824578      -.412205
      7   1.200000E+01   2.769677E+01  -1.569677E+01     -1.671324         1.000000      10.606269     -1.479952
      8   2.700000E+01   1.500135E+01   1.199865E+01      1.277565         1.000000      10.917972      1.098982
      9   4.800000E+01   3.804862E+01   9.951384E+00      1.059580         1.000000      11.269045       .883073
     10   5.000000E+01   4.890666E+01   1.093337E+00       .116414         1.000000      10.065968       .108617


          LISTING OF THE DATA ANALYZED

   1          2.50000    25.00000    25.00000     1.50000    34.00000    64.00000

   2         13.00000    21.00000    21.00000      .87000    36.00000    65.00000

   3          3.50000    22.00000    22.00000      .43000    41.00000    82.00000

   4          1.75000     9.00000     1.30000     1.80000    15.00000    23.00000

   5          3.00000    23.00000    23.00000     2.00000    33.00000    64.00000

   6          2.00000    10.00000      .60000     3.30000    13.00000    16.00000

   7          5.50000     7.00000     1.40000     3.40000    16.00000    12.00000

   8          6.00000     6.00000      .80000     5.00000    11.00000    27.00000

   9          1.30000     8.00000     2.70000     1.50000    19.00000    48.00000

  10          5.00000    18.00000     3.60000     1.80000    27.00000    50.00000
```

COST ESTIMATE

For the job listed on the Sample Output, the central processor unit time was 0.612 seconds. At the current rate for the University of Minnesota ($0.20/sec.), the computer time cost $0.12 plus a small charge for output supplies.

Charge to user = computer costs + postage + network overhead
= $0.30 + postage + network overhead

CONTENTS—UMST580

pages
1- 3 Identification & Abstract
5–11 User Instructions
13–16 I/O
17 Cost—Contents

DESCRIPTIVE TITLE	Cross Classification
CALLING NAME	UMST590
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Robert Ellis, Department of Journalism University of South Dakota
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program generates a cross-tabulation, i.e. a bivariate frequency distribution, for pairs of variables. As an option, a distribution may be formed for each value of a third variable or for each pair of values of a third and a fourth variable, i.e., using one or two control variables. Missing data options may interpret blanks and non-numerics as missing data, or read missing data tags, or convert plus and minus punches to numeric values. The data may be used as given, or they may be grouped into intervals before a plot is formed. The data may be real or integer numbers. A transgeneration subprogram is included.

The following may be selected as output.

1. Chi square and contingency coefficient C.
2. Column and row percentages and percentages to the overall total of the table.
3. Nonparametric correlation coefficients: Kendall tau with associated z; Goodman—Kruskal gamma, and Somers dyx and dxy. (Note: for a 2X2 table, gamma = Yule's Q.)

Limitations

1. Number of variables $2 \leq NVAR \leq 80$
   Number of observations $NOBS \leq 40,000$

continued

2. Total data input (NOBS X NV) $\leq$ 40,000 where NV is the largest integer contained in $\frac{\overline{N}\text{VAR} + 9}{10}$;
   e.g., 10,000 observations on 40 variables, or
   5,000 observations on 80 variables.
3. The number of levels on any variable must be $\leq$ 20, not including missing data. Values taken on may be any real number, but only 20 distinct values may occur. In order to meet this restriction, Maximum and Minimum Cards, transgeneration or one of the recoding options (Interval or PICK) may be used.
4. Maximum table size is 20 X 20.

Formulae

Let x_{ij} be the data element in the ith row and jth column.

Sums of columns:

$$n_{.j} = \sum_{i=1}^{r} X_{ij} \text{ where r is the number of rows.}$$

Sums of rows:

$$n_{i.} = \sum_{j=1}^{c} X_{ij} \text{ where c is the number of columns.}$$

Total:

$$N = \sum_{j=1}^{c} n_{.j}$$

Chi square:

$$\text{Chi square} = \sum_{i=1}^{r} \sum_{j=1}^{c} \left[\frac{[X_{ij} - \frac{n_{.j}n_{i.}}{N}]^2}{\frac{n_{.j}n_{i.}}{N}} \right] = N\left[\left\langle \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{X_{ij}^{2}}{n_{i.}n_{.j}} \right\rangle - 1 \right]$$

The latter form is used in the program.

Degrees of freedom: d.f. = (r–1)(c–1)
Where missing data or subtables result in zero rows or columns, the message 'ADJUSTED TO ______ FOR THIS TABLE' will appear.

Yates' continuity correction, applied to 2 X 2 tables:

If $X_{ij} < f_{ij}$ then use $X_{ij} + 1/2$

continued

If $X_{ij} > f_{ij}$ then use $X_{ij} - 1/2$

If $X_{ij} = f_{ij}$ then use X_{ij}

Expected frequencies:

$$f_{ij} = n_{i.} \text{ X } n_{.j}/N$$

Contingency coefficient:

$$C = \sqrt{\text{Chi square} / (N + \text{Chi square})}$$

where Chi square has not been corrected for continuity in the case of a 2 X 2 table.

$$Q = \sum_{i=1}^{r-1} \sum_{j=2}^{c} (X_{ij} X_{kl}) \quad \text{where } k = (i+1), \ldots, r \text{ and } l = 1, \ldots, (j-1)$$

$$P = \sum_{i=1}^{r-1} \sum_{j=1}^{c-1} (X_{ij} X_{kl}) \quad \text{where } k = (i+1), \ldots, r \text{ and } l = (j+1), \ldots, c$$

$$S = P - Q$$

Kendall Tau:

1. For square table (i.e., r = c):

$$\text{Tau B} = \frac{S}{\sqrt{1/2N(N-1)-T} \quad \sqrt{1/2N(N-1)-U}}$$

$$\text{where } T = \frac{1}{2} \sum_{i=1}^{r} n_{i.}(n_{i.} - 1)$$

$$\text{and } U = \frac{1}{2} \sum_{j=1}^{c} n_{.j}(n_{.j} - 1)$$

2. For non-square table (i.e., r ≠ c):

$$\text{Tau C} = \frac{2S}{N^2(\frac{m-1}{m})} \quad \text{where m is smaller of r, c}$$

continued

Sampling variance of S:

$$\mathrm{Var} = \frac{1}{18}\left[N(N-1)(2N+S) - \sum_{i=1}^{r} n_{i.}(n_{i.}-1)(2n_{i.}+5) - \sum_{j=1}^{c} n_{.j}(n_{.j}-1)(2n_{.j}+5)\right]$$

$$+ \frac{1}{9N(N-1)(N-2)}\left[\sum_{i=1}^{r} n_{i.}(n_{i.}-1)(n_{i.}-2)\right]$$

$$\left[\sum_{j=1}^{c} n_{.j}(n_{.j}-1)(n_{.j}-2)\right] + \frac{1}{2N(N-1)}\left[\sum_{i=1}^{r} n_{i.}(n_{i.}-1)\right]$$

$$\left[\sum_{j=1}^{c} n_{.j}(n_{.j}-1)\right]$$

Normal deviate:

$$Z(s) = \frac{S}{\sqrt{\mathrm{Var}}}$$

(Continuity correction): Given a 2 X 2 table the absolute value of S is reduced by 1/2 N with the restriction that if $|S| \leq 1/2\,N$, $S = 0$. For all other tables, $|S|$ is reduced by 1.

Goodman-Kruskal Gamma:

$$\mathrm{Gamma} = S/(P+Q)$$

Somers d:

$$d_{xy} = \frac{S}{\frac{1}{2}\left[N^2 - \left(\sum_{i=1}^{r} n_{i.}^2\right)\right]} \qquad d_{xy} = \frac{S}{\frac{1}{2}\left[N^2 - \left(\sum_{j=1}^{c} n_{.j}^2\right)\right]}$$

continued

REFERENCES

Hays, W.L., *Statistics for Psychologists* (New York: Holt, Rinehart & Winston Inc., 1963), pp. 652-656.

Kendall, M.G., *Rank Correlation Methods* (3rd ed.; New York: Hafner Pub. Co., 1962).

McNemar, Q., *Psychological Statistics* (3rd ed.; New York: John Wiley & Sons, Inc., 1962).

Somers, R.H., "A New Asymmetric Measure of Association for Ordinal Variables," Am. Soc. Rev., *27* (1962), pp. 799–816.

USER INSTRUCTIONS

System Control Cards

System Control Cards will be prepared by Univ. of Minnesota personnel.

Program Control Cards

Problem Card

* indicates information which must be provided. Other fields are options and may be left blank if not desired.

Columns	*Parameter*	*Contents*
*1- 8		PROBLEM
*8–12	NOBS	Positive integer: the number of observations or sets of Data Cards to be read in. 0 or blank: number of observations is counted by computer. An I field must be provided for counting and an End of Data Card must follow the data (see Optional Cards). The I field must be the first field read and is not counted as a variable.
*13–14	NVAR	Number of variables (does not include the optional counting variable; NVAR $\le$ 80).
*15–16	NFC	Number of Variable Format Cards. If zero or blank, NFC = 1 is assumed. 0 $\le$ NFC $\le$ 9.
17–18	NTGC	Number of Transgeneration Cards (optional).
19–20	NVA	Number of variables added by transgenerations (optional).
21–22	NLV	Number of labeled variables (optional).
*23–24	MODE	1: input of NVAR variables (no missing data). Format Card must read data with X, F or E fields. 2: input of NVAR variables. All blank and all non-numerical fields are considered as missing data.

continued

11

Columns	*Parameter*	*Contents*
*23–24		Format Cards must read data with X or A fields. Data must be integers only. 3: input of NVAR variables, first reading a Missing Data Tag Card (see Optional Cards). Then identity between Tag Card and data indicates missing data. Blanks are read as zeros. Format Card must read data with X, F or E fields. 4: input of NVAR variables. A minus punch is converted to a value of 11, a plus punch to a value 12. Data must be read with X or A format fields. 5: same as mode = 4 except that in addition a 0 is converted to a value of 10.
25–26	NIC	Number of Interval Cards to be read. NIC ≤ NVAR + NVA (optional).
27–28	NPC	Number of Pick Cards to be read, NPC ≤ NVAR + NVA (optional).
*29–30	NSC	Number of Select Cards to be read. NSC ≤ 99.
31	MAP	1: table output is suppressed. 0 or blank: table is printed.
32	NCHI	1: Chi square and contingency coefficient printed. 2: expected Frequencies and Chi square contingency printed. 0 or blank: otherwise.
33	NPCT	1: table of percentages to overall total is desired. 0 or blank: otherwise.
34	NPCTR	1: percentages of row total are desired. 0 or blank: otherwise.
35	NPCTC	1: percentages of column total are desired. 0 or blank: otherwise.
36	NTAU	1: nonparametric correlation coefficients are desired. 0 or blank: otherwise.

continued

Columns	*Parameter*	*Contents*
37	MAXMIN	1: Maximum and Minimum Value Cards are to be read. (see Optional Cards) 0 or blank: otherwise
38	LTAPE	1: data tape is to be rewound before the data is read. This must not be used if data tape is on cards or it will destroy the computer program for that run. 0: data tape is not to be rewound by the program.
39	KORD	0 or blank: tables will be printed so that values on the vertical axis run from high at top to low at bottom (axes in same directions as a correlation diagram). 1: values on vertical axis are reversed in order so low values are at top, high at bottom.
40		blank
41–80	IDENT	Alphanumeric identification of the problem.

Format Cards

The Format Cards must provide for reading one value for every variable one observation at a time. The variables are considered to be indexed or numbered by a subscript assigned according to the order in which they are read from the Data Cards. The Format Cards must contain only X, E, or F fields unless an I field is needed for counting (see Optional Cards), or the missing data mode requires A fields to be used.

Data Input

The Data Cards must be punched in accordance with the format given on the Format Card. One observation must be read for each variable from one card or set of cards. If NOBS is not given on the program, there must be a non-negative integer in the first field (an I field) read by the Format Card for use in counting observations. (See also End of Data Cards).

Select Cards

These cards specify table construction. The number of cards must be punched in Cols. 29–30 of the Problem Card. The maximum

continued

table size is 20 X 20. Note that the amount of computer time used is directly proportional to the number of Select Cards.

Columns	*Contents*
1– 6	SELECT
7– 9	The index number of the dependent[α] (row) variable. I.e., the order number corresponding to the order in which that variable is read from the Data Card.
10–12	Number of the first control variable, if any.
13–15	Number of the second control variable, if any.
16–18	The number of independent variables to be read from the following list, $\leq$ 30.
19–20	blank
21–22	The number of the first independent variable.
23–24	The number of the second independent variable.
⋮	⋮
79–80	The number of the thirtieth independent variable.

[α] This variable will be referred to as the dependent variable although it need not be such to the user. It will always appear on the vertical axis of the table.

Note: A bivariate distribution will be formed between the dependent variable and each of the independent variables in turn. If a variable is indicated for the first control variable a distribution will be formed for each level of the first control variable, e.g., if the first control variable has five levels, five tables will be formed between each independent variable and the dependent variable. Each of these sub-tables will be based only on observations having the same value on the control variable. If two control variables are specified, a sub-table will be formed for each combination of levels on the two controlled variables, e.g., if the first control variable has five levels and the second control variable has four, then twenty sub-tables will be formed.

Caution: The ingenious user will find it theoretically possible to generate 1.2 million tables (and 2.4 million pages of output) for each Problem Card. Therefore, considerable self-restraint is required in the selection of tables and especially sub-tables.

continued

Output of percentages should be requested only if necessary, and then only that one set (by row, column, or entire table) which will be most useful. As a general rule, no more than 100 to 200 tables (depending on tables size and output options) should be requested by any one job. Users who insist on requesting large numbers of tables should do one or all of the following:

a. do preliminary analysis with a straight correlation program;
b. suppress table printout of this program by placing a 1 in Col. 31 of the Problem Card thereby obtaining statistics only;
c. advise the computer operator by a prominent note attached to the job deck how many pages of output are expected, and be prepared to wait.

Finish Card

Columns	*Contents*
1– 6	FINISH

This card terminates the program.

Optional Cards

Transgeneration Cards

Each Transgeneration Card specifies the transformation or variable generation desired. The number of Transgeneration Cards must be specified in Cols. 17–18 of the Problem Card. The new variable or transformed variable may replace one of the existing variables or may be added to the original set of variables. Transformations are done sequentially; thus, two variables may be added in one transformation and the square root of the sum may be done in the next transformation. The number of variables added (NVA) to the original set must be specified in Cols. 19–20 of the Problem Card. There are NVAR variables originally read into the problem from input cards, to which NVA variables are added, resulting in NFVAR final variables. NFVAR must be $\leq$ 80. Any of the NFVAR variables can be changed or relocated as desired by the appropriate indication on a Transgeneration Card. At no point in the program may an index number larger than NFVAR be used. This option handles missing data; the result of the transgeneration will be a missing data code whenever any of the elements used in the transgeneration for a particular observation is missing, e.g., if variables x and y are to be added together, the result for observation n will be a missing

continued

data code if either x or y (or both) is missing for that observation. If many variables are to be summed and only the sum used thereafter, it is recommended that the user do this in a separate program and use the resulting data file (containing only those variables to be cross-classified) as input to this program.

Notation

x:	the variable prior to transgeneration
x':	the transformed variable which (depending on the instruction in the Transgeneration Card) replaces the original variable x or becomes a new variable
c:	a constant
A,B:	specified variable numbers, or indices
*:	indicates multiplication

Code	*Transgeneration*	*Restrictions*
01	$x' = \sqrt{x}$	$x \geq 0$
02	$x' = \sqrt{x} + \sqrt{x+1}$	$x \geq 0$
03	$x' = \log_{10}(x)$	$x > 0$
04	$x' = e^x$	$x < 740$
05	$x' = \arcsin \sqrt{x}$	$-1 \leq x \leq 1$
06	not defined	
07	$x' = 1/x$	$x \neq 0$
08	$x' = x + c$	
09	$x' = x*c$	
10	$x' = x^c$	$x > 0$
11	$x' = x_A + x_B$	
12	$x' = x_A - x_B$	
13	$x' = x_A*x_B$	
14	$x' = x_A/x_B$	$x_B \neq 0$

continued

Code	*Transgeneration*	*Restriction*
15	if $x \leq c$, set $x' = 1$ otherwise, $x' = 0$	
16	if $x_A \geq x_B$, set $x' = 1$ otherwise, $x' = 0$	
17	$x' = \log_e(x)$	$x > 0$
18	not defined	
19	not defined	
20	$x' = \sin(x)$	$\lvert x\rvert < 2.2 \times 10^{14}$, x in radians
21	$x' = \cos(x)$	$\lvert x\rvert < 2.2 \times 10^{14}$, x in radians
22	$x' = \arctan(x)$	
23	$x' = {x_A}^{x_B}$	$x_A > 0$
24	$x' = c^x$	$c > 0$
25	if $x = c$, set x' to missing data code	
26	if $x \geq c$, set x' to missing data code	
27	if $x < c$, set x' to missing data code	
28*	if $x < 0$, set $x' = 11$ if $x > 0$, set $x' = 12$ otherwise, $x' = x$	
29*	if $x = 0$, set $x' = 10$ if $x < 0$, set $x' = 11$ if $x > 0$, set $x' = 12$ otherwise, $x' = x$	

**Note*: Codes 28 and 29 simulate Mode 4 and Mode 5 for particular variables. These variables must be read in an A format with Mode = 1 and all other variables must be read in an F format. Non-numerics other than those specified will be treated as missing data. Maximum and Minimum Value Cards may not be used. Pick Cards may not be used with these variables. These Transgeneration Cards must preceed any others involving the same

continued

variables. Interval Cards may be used. The use of these codes is not recommended unless absolutely necessary, and then the user should see a consultant familiar with the program.

Columns	*Contents*
1– 6	TRNGEN
7– 9	Variable number to be assigned to x' ($\leq$ NVAR + NVA)
10–11	Transgeneration code (01...29)
12–14	Variable index number A, or index number of x
15–20	Variable index number B, or constant c (punched right-justified in Col. 20, unless decimal point is punched)

Label Cards

Identification labels may be printed as part of the table for any or all of the variables. The number of labels to be read, if any, must be punched in Cols. 21–22 of the Problem Card. The labels are punched one to seven to a card in accordance with the format below and may be in any order; any of the five fields on the card may be used or left blank so long as the other labels and corresponding variable numbers are in the correct column. Variables for which a label is not provided will be labelled with blanks.

Columns	*Contents*
1– 5	LABEL
6– 8	blank
9–10	Variable number for following label.
11–20	Any 1–10 character identification for the preceding variable number
21–23	blank
24–25	Another variable number
26–35	Corresponding identification
36–38	blank
39–40	Variable number
41–50	Corresponding identification
51–53	blank

continued

Columns	*Contents*
54–55	Variable number
56–65	Corresponding identification
66–68	blank
69–70	Variable number
71–80	Corresponding identification

Interval Cards

These cards instruct the computer to collapse data into intervals as specified on the cards and set maximum and minimum values. One card is required for each variable to be grouped. The number of such cards must be punched in Cols. 25–26 of the Problem Card. The upper limits for each interval to be formed are specified on the card, except that any value less than the smallest value or greater than that largest value will be treated as missing data. The upper limit of the first interval will be the second smallest value. The values on the card may be in any order. The values are read with a format of F5.0, i.e., the decimal point is assumed to be at the right unless it is punched.

Columns	*Contents*
1– 6	INTRVL
7– 8	Variable number
9–10	Number of interval values to be read ($\leq$ 21), i.e., the number of intervals plus one.
11–15	The minimum value, below which any data are considered as missing.
16–20	Largest value to be included in the first (lowest) interval.
21–25	Largest value to be included in second interval.
⋮	Continue in 5–columns fields for the number of values punched in Cols. 9–10, beginning in Col. 1 of second card if necessary. End with largest value to be included in the highest interval. Higher values will be interpreted as missing data.

Pick Cards

These cards provide a means for common data-reordering problems which otherwise would be quite difficult with standard transgenerations. The card contains 20 fields corresponding to data

continued

input values of 0...19. The number punched by the user is the value to be given to the corresponding input value. For example, if the user wishes to recode all 0's, 2's, and 4's to a value of 1, and all 1's, 3's, and 5's to a value of 2, then the number "1" should be punched in fields 0, 2 and 4, and the "2" in fields 1, 3, and 5. Input values for which a number is not supplied by the user will be treated as missing data. The number of cards of this type must be punched in Cols. 27–28 of the Problem Card.
Note: Input values less than 0 or greater than 19 will be treated as missing data. Floating point numbers within the above range will be truncated to integers before recoding, e.g., 5.307 will be truncated to 5.

Columns	*Contents*
1– 4	PICK
5– 6	blank
7– 8	Variable number
9–10	blank
11–12	Value to be given to an input value of 0 (Caution: If field is not used, leave blank, do not fill with 00's unless conversion to 0 is desired.)
13–14	Value to be given to an input value of 1
15–16	Value to be given to an input value of 2
⋮	⋮
49–50	Value to be given to an input value of 19
51–80	blank

PICK and INTRVL options may not be used for the same variables.

Missing Data Identification Card(s)

This card is required if MODE = 3 in Cols. 23–24 of the Problem Card. It is punched exactly the same as the Data Cards and is read from File 2 by the input format specified by the user. One value is read off this card for each variable. Each Data Card set is matched with the Missing Identification set to identify missing values. If the value of a variable on a Data Card set is the same as the value for that variable on the Missing Data Identification set, the value of the variable on the Data Card is identified as missing.

continued

Identification of an input data value for observation n of variable x as missing results in observation n being dropped for all distributions (and subsequent statistics) involving variable x but the observation is retained for distributions not involving variable x.

Maximum and Minimum Value Cards

These cards are required if Col. 37 of the Problem Card = 1. One value is read off this card for each variable according to the data input format specified by the user. The cards are punched in exactly the same format as the Data Cards. Each Data Card set is matched with the Maximum Value set and the Minimum Value set to identify missing (out of range) data. If the value of a variable in the data set is less than the value for that variable on the Minimum Value set or greater than the value of that variable on the Maximum Value set, the value of the variable on the Data Card will be identified as missing. The Maximum Value set is read first, then the Minimum Value set. If the Maximum Value for a given variable is less than the Minimum Value for the same variable, an error message will be issued and the job aborted. If the maximum and minimum values for the same variable are equal (including both blank), the maximum will be set to 9.0×10^{100} and the minimum value to -9.0×10^{100}.

End of Data Cards (This card is used only if NOBS is 0 or blank) It must have a –1 punched in the first (I) field as specified by the Format Card. If the Format Card specified more than one Data Card read as a set, the –1 card must be followed by the appropriate number of blank cards, to make a complete set.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

Program Control Cards for problem 1
  Problem Card
  Format Card(s)
  Transgeneration Cards (optional)
  Label Cards (optional)
  Interval Cards (optional)
  Pick Cards (optional)
  Missing Data Value Cards (optional)
  Maximum Value Cards (optional)
  Minimum Value Cards (optional)

continued

Data Cards for problem 1 (unless separate data file is used)
End of Data Card for problem 1 (optional)
Select Cards for problem 1
Program Control Cards for problem 2
Data Cards for problem 2
End of Data Card for problem 2
Select Cards for problem 2
:
:
Finish Card

Processing Order

The use of several program options in the same problem may result in complex interactions among them, and, if care is not excerised, unwanted results. The program uses the following processing order in converting data:

1. Any conversion specified by Mode (including detection of blanks and non-numerics or missing data value comparison, as opted).
2. Check against maximum and minimum values (if opted).
3. Recoding specified by PICK Card (if opted).
4. Conversion by TRNGEN Cards (if opted).
5. Conversion by INTRVL Cards (if opted).
6. A frequency distribution of current values is formed for each variable. At this point the number of different values for a given variable (excluding missing data) must not exceed 20; if it does, an error message will be issued and the job aborted.
7. Construction of tables according to SELECT Cards.

Note 1: A missing identification code (7HMISSING), once generated, will cause the element to be by-passed in subsequent transgenerations or other data conversions.
Note 2: PICK and INTRVL options are mutually exclusive for a given variable.

REFERENCE

1. Anderson, D., and Frisch, M., *UMST Computer Programs Manual: Statistical Programs for Use on the Control Data Corporation 6600 Computer* (Revised; Minneapolis, Minn.: Univ. of Minn., Univ. Comp. Ctr., Fall 1969), Introduction, Sec. VII.

SAMPLE INPUT

```
PROBLEM       2 1       1     1 21111    KRAM TEST NEXT GOOD PROBLEM
(I2,8X,2F1)
          39
          11
          11
          31
          31
          19
          39
          51
          21
          23
          51
          10
          29
          31
          32
          32
           1
          11
          11
          11
          31
          51
-1
SELECT  1        1   2
FINISH
```

SAMPLE OUTPUT

```
UMST590 (SSRFC) CROSS CLASSIFICATION                    KRAM TEST NEXT GOOD PROBLEM

        NOBS =     -0   NVAR =  2   NFC =  1   NTGC = -0   NVA = -0   NLV = -0
        MODE =  1   NIC = -0   NPC = -0   NSC =  1   OPTIONS = -0 2 1 1 1 1-0-0-0

        INPJT FORMAT IS (I2,8X,2F1)


        NO. OF OBSERVATIONS READ =    22

        THIS PROBLEM REQUIRES A FIFLD LENGTH OF 00043300
```

continued

```
          ROW VARIABLE  1               BY COL VARIABLE  2

VALUES          TOTAL          0       1       2       3       9


   5,000          3            0       3       0       0       0
   3,000          8            0       4       2       0       2
   2,000          3            0       1       0       1       1
   1,000          7            1       5       0       0       1
 -0.              1            0       1       0       0       0

       TOTAL     22            1      14       2       1       4


          EXPECTED FREQUENCIES (ROUNDED)

   5,000                      .1     1.9      .3      .1      .5
   3,000                      .4     5.1      .7      .4     1.5
   2,000                      .1     1.9      .3      .1      .5
   1,000                      .3     4.5      .6      .3     1.3
 -0.                          .0      .6      .1      .0      .2
          CHI SQUARE =      14.910     C =        .636      DF =     16

          TAU-B =             .098     Z(S) =     .484      S =        1.500E+01
          GAMMA =             .152     DYX =      .113      DXY =       .085


          PERCENTAGE OF TOTAL

   5,000                     0.0    13.6     0.0     0.0     0.0
   3,000                     0.0    19.2     9.1     0.0     9.1
   2,000                     0.0     4.5     0.0     4.5     4.5
   1,000                     4.5    22.7     0.0     0.0     4.5
 -0.                         0.0     4.5     0.0     0.0     0.0

       TOTAL     22


          PERCENTAGE OF COLUMN TOTAL

   5.000                     0.0    21.4     0.0     0.0     0.0
   3.000                     0.0    28.6   100.0     0.0    50.0
   2.000                     0.0     7.1     0.0   100.0    25.0
   1.000                   100.0    35.7     0.0     0.0    25.0
 -0.                         0.0     7.1     0.0     0.0     0.0

       TOTAL                   1      14       2       1       4


          PERCENTAGE OF ROW TOTAL

   5,000          3          0.0   100.0     0.0     0.0     0.0
   3,000          8          0.0    50.0    25.0     0.0    25.0
   2,000          3          0.0    33.3     0.0    33.3    33.3
   1.000          7         14.3    71.4     0.0     0.0    14.3
 -0.              1          0.0   100.0     0.0     0.0     0.0
```

COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 0.306 seconds. At the current rate for the Univ. of Minnesota ($0.20/sec.), the computer time cost $0.06 plus a charge for output supplies.

Charge to user = computer costs + postage + network overhead
= $0.16 + postage + network overhead

CONTENTS—UMST590

pages
1– 5 Identification & Abstract
7–18 User Instructions
19–21 I/O
23 Cost—Contents

DESCRIPTIVE TITLE	Descriptive Statistics
CALLING NAME	UMST600
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Decks and listings currently available
CONTACT	William Craig, EIN Tech. Rep. Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

UMST600 provides a nearly complete description of the distribution of each of up to 999 variables. For each variable (X) the following measures are calculated: $\sum X$, $\sum X^2$, $\sum(X-\overline{X})^2$, mean, variance, and standard deviation. Each measure is produced after reading in all observations, but options exist to identify missing data so that the measures may pertain to a different subset of the total number of observations for each variable. The count of non-missing observations is output for each variable. There is no limit on the number of observations to be processed.

Several options are available which make the program more useful. The user may opt to produce $\sum X^3$, $\sum X^4$, $\sum(X-\overline{X})^3$ and $\sum(X-\overline{X})^4$. He may also opt to print the maximum and minimum value for each variable. X value (from largest to smallest), frequency, cumulative frequency, cumulative proportion and absolute proportion for each variable may be output. Choosing this option limits the size of the job that can be processed in one use of the program to 400 variables and 100 observations or 10 variables and 4000 observations. Variables may also be named, attaching an eight-character label to each variable's output.

continued

Notation and Computations

N = the number of observations

Variance formula 1

$$s^2 = \frac{N \Sigma X^2 - (\Sigma X)^2}{N(N-1)} \quad \text{(unbiased)}$$

Variance formula 2

$$s^2 = \frac{N \Sigma X^2 - (\Sigma X)^2}{N^2} \quad \text{(biased)}$$

$$\Sigma(X - \overline{X})^2 = \Sigma X^2 - N\overline{X}^2$$

$$\Sigma(X - \overline{X})^3 = \Sigma X^3 - 3\overline{X}\Sigma X^2 + 2N\overline{X}^3$$

$$\Sigma(X - \overline{X})^4 = \Sigma X^4 - 4\overline{X}\Sigma X^3 + 6\overline{X}^2\ \Sigma X^2 - 3N\overline{X}^4$$

USER INSTRUCTIONS

System Control Cards

System Control Cards will be prepared by Univ. of Minnesota personnel.

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8–10	NVAR	Number of variables to be read in ($1 \leq \text{NVAR} \leq 999$; does not include the optional counting variable)
11–14	NOBS	Number of Data Cards or sets of Data Cards to be read in. If blank or zero, computer will count observations and an End of Data Card must be supplied. See End of Data Cards. ($0 \leq \text{NOBS} \leq 9999$)
15		blank
16	NFC	Number of Format Cards ($\text{NFC} \leq 9$) If zero or blank, NFC = 1 is assumed.
17–18	NN	1: Name Cards to be read in 0: no Name Cards used
19–20	MODE	1: read NVAR variables, with F or X fields. No missing data. 2: read NVAR variables with A or X fields. All blanks and all non-numeric fields are considered as missing data. Data must be integers. 3: read NVAR variables with F or X fields. Data must be preceded by Missing Data Identification Cards. See below.
21	N1	1: output third and fourth moments 0: otherwise
22	N2	1: output maximum and minimum values 0: otherwise
23	N3	1: output frequency distribution 0: otherwise

continued

Columns	*Parameter*	*Contents*
24	BIAS	1: use formula 1, unbiased estimate, for variance 0: use formula 2, biased estimate, for variance
25–80	IDENT	Alphanumeric identification of the problem

Format Card(s)

Variables are considered to be indexed, or numbered, by a subscript according to the order in which they are read in from the Data Cards. Data Cards are read from the first file (default name TAPE2). The first variable read is No. 1, the second is No. 2, etc., and they are denoted by X_1, X_2, ..., X_J. The Format Card must provide for reading in the J elements of an observation at one time in F or X fields, unless MODE = 2 is used, in which case all data are in A fields, of 10 columns or less. In addition, if Cols. 11–14 of the Problem Card are blank, the first field read by the Format Card shall be an integer field of at least 2 columns. This field is used in counting observations.

Name Card(s) (Necessary if NN=1, Cols. 17–18, Problem Card)

An alphabetic identification of up to 8 characters may be given to each variable by reading in Name Cards.

Columns	*Contents*
1– 8	Name for variable 1
9–16	Name for variable 2
⋮	⋮
73–80	Name for variable 10

This format is repeated with 10 names to a card until each of the NVAR variables has been assigned a name.

Missing Data Identification Card(s) (Necessary for input MODE = 3)

A Missing Data Identification Card is a card or set of cards which has a value punched for every variable to be read. This value must be the missing data value; i.e., a value which is not an actual data value. All missing pieces of data must be punched with this value. The value may be different for each variable. Blanks in the Data Cards are read as zeros and

continued

considered as data unless zero is the missing value. Missing Data Identification Cards are read by the same format as supplied on the Format Cards.

Input Data Cards

The Data Cards must be punched in accordance with the format given on the Format Card. One value for each variable is read at a time from a set of Data Cards. Each set of values starts on a new set of cards. If Cols. 11–14 of the Problem Card are blank, there must be a non-negative integer in the first field (an I field) read by the Format Card, for use in counting observations.

End of Data Card(s)

This card (or set of cards) is required only if Cols. 11–14 of the Problem Card are 0 or blank. It must have –1 punched where it will be read by the first (i.e. integer) field of the Format Card. If the Format Card indicates reading data from more than one card, the –1 card must be followed by enough blank cards to form a data set.

Finish Card

Columns	*Contents*
1– 6	FINISH

Used after last problem.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

Program Control Cards for problem 1
    Problem Card for problem 1
    Format Card(s) for problem 1
    Name Card(s) for problem 1 (if any)
    Missing Data Identification Card(s) for problem 1
      (if required)
Data Cards for problem 1
End of Data Card(s) for problem 1 (if required)
Program Control Cards for problem 2
Data Cards for problem 2

continued

End of Data Card(s) for problem 2 (if required)
:
Program Control Cards for last problem
Data Cards for last problem
End of Data Card(s) for last problem (if required)
Finish Card

Note: It is important not to request unnecessary options for output as this may generate huge quantities of paper.

SAMPLE INPUT

```
PROBLEM  6     1 1 10001 UMST600--BASIC OUTPUT
(I2,2X,F2.0,2X,F3.0,4F4.0)
AGE     IQ      SPELLINGARITH   ENGLISH READING
01  10  100  16  98    8 5.0
02  10  104  16  98    8 5.5
03  10  104  20  72    7 6.5
04  12   94  14  90    8 5.0
05  12  100  20  89    7 5.5
06  10  102  17  92    7 5.0
-1
PROBLEM  4   6 1 0 11111 UMSTAT60--FREQUENCY OUTPUT
(11X,4F4.0)
01  10  100  16  98    8 5.0
02  10  104  16  98    8 5.5
03  10  104  20  72    7 6.5
04  12   94  14  90    8 5.0
05  12  100  20  89    7 5.5
06  10  102  17  92    7 5.0
FINISH
```

000 0146

SAMPLE OUTPUT

```
UMST600                DESCRIPTIVE STATISTICS  UMST600--BASIC OUTPUT

       6 VARIABLES      -0 OBSERVATIONS     1 FORMAT CARDS
      NAME CARD OPTION     1 = YES   MODE =  1
      3RD AND 4TH MOMENTS  OPTION    0 = NO
      MAX AND MIN VALUE    OPTION    0 = NO
      FREQUENCY DISTRIB    OPTION    0 = NO
      UNBIASED VARIANCE    OPTION    1 = YES

     FORMAT = (I2.2X.F2.0.2X.F3.0.4F4.0)

     THIS PROBLEM REQUIRES A FIELD LENGTH OF 030400B

       6 OBSERVATIONS COUNTED

VARIABLE     NOBS       MEAN         ST. DEV.      VARIANCE        SUM(X)            SUM(X2)           SUM(X-XB)2

 1 AGE         6    1.066667E+01  1.032796E+00  1.066667E+00    6.400000000E+01  6.880000000E+02  5.333333333E+00

 2 IQ          6    1.006667E+02  3.723797E+00  1.386667E+01    6.040000000E+02  6.087200000E+04  6.933333333E+01

 3 SPELLING    6    1.716667E+01  2.401388E+00  5.766667E+00    1.030000000E+02  1.797000000E+03  2.883333333E+01

 4 ARITH       6    8.983333E+01  9.558591E+00  9.136667E+01    5.390000000E+02  4.887700000E+04  4.568333333E+02

 5 ENGLISH     6    7.500000E+00  5.477226E-01  3.000000E-01    4.500000000E+01  3.390000000E+02  1.500000000E+00

 6 READING     6    5.416667E+00  5.845226E-01  3.416667E-01    3.250000000E+01  1.777500000E+02  1.708333333E+00
```

continued

```
UMST600                    DESCRIPTIVE STATISTICS  UMSTAT60--FREQUENCY OUTPUT

          4 VARIABLES       6 OBSERVATIONS     1 FORMAT CARDS
         NAME CARD OPTION     0 = NO     MODE =  1
         3RD AND 4TH MOMENTS  OPTION    1 = YES
         MAX AND MIN VALUE    OPTION    1 = YES
         FREQUENCY DISTRIB    OPTION    1 = YES
         UNBIASED VARIANCE    OPTION    1 = YES

        FORMAT = (11X,4F4.0)

        THIS PROBLEM REQUIRES A FIFLD LENGTH OF 030400B

VARIABLE     NOBS        MEAN       ST. DEV.     VARIANCE        SUM(X)          SUM(X2)        SUM(X-XB)2

 1            6     1.716667E+01  2.401388E+00  5.766667E+00   1.030000000E+02  1.797000000E+03  2.883333333E+01

 2            6     8.983333E+01  9.558591E+00  9.136667E+01   5.390000000E+02  4.887700000E+04  4.568333333E+02

 3            6     7.500000E+00  5.477226E-01  3.000000E-01   4.500000000E+01  3.390000000E+02  1.500000000E+00

 4            6     5.416667E+00  5.845226E-01  3.416667E-01   3.250000000E+01  1.777500000E+02  1.708333333E+00


VARIABLE        SUM(X3)           SUM(X4)          SUM(X-XB)3       SUM(X-XB)4        MAX. VALUF        MIN. VALUE

 1         3.1849000E+04     5.7300900E+05     1.0555556E+01     2.3315278E+02      2.00000E+01       1.40000E+01

 2         4.4682890E+06     4.1133902E+08    -4.5725556E+03     1.1006049E+05      9.80000E+01       7.20000E+01

 3         2.5650000E+03     1.9491000E+04     0.                3.7500000E-01      8.00000E+00       7.00000E+00

 4         9.8237500E+02     5.4901875E+03     1.0555556E+00     1.4678819E+00      6.50000E+00       5.00000E+00
       FREQUENCY TABLE FOR VARIABLE  1

            VALUE             FREQ          CUM FREQ       CUM PROP       ABS PROP

        2.0000000E+01           2              6           1.0000          .3333
        1.7000000E+01           1              4            .6667          .1667
        1.6000000E+01           2              3            .5000          .3333
        1.4000000E+01           1              1            .1667          .1667
       FREQUENCY TABLE FOR VARIABLE  2

            VALUE             FREQ          CUM FREQ       CUM PROP       ABS PROP

        9.8000000E+01           2              6           1.0000          .3333
        9.2000000E+01           1              4            .6667          .1667
        9.0000000E+01           1              3            .5000          .1667
        8.9000000E+01           1              2            .3333          .1667
        7.2000000E+01           1              1            .1667          .1667
       FREQUENCY TABLE FOR VARIABLE  3

            VALUE             FREQ          CUM FREQ       CUM PROP       ABS PROP

        8.0000000E+00           3              6           1.0000          .5000
        7.0000000E+00           3              3            .5000          .5000
       FREQUENCY TABLE FOR VARIABLE  4

            VALUE             FREQ          CUM FREQ       CUM PROP       ABS PROP

        6.5000000E+00           1              6           1.0000          .1667
        5.5000000E+00           2              5            .8333          .3333
        5.0000000E+00           3              3            .5000          .5000
```

COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 0.221 seconds. At the current rate for the Univ. of Minnesota ($0.20/sec.), the computer time cost $0.04 plus a charge for output supplies.

Charge to user = computer costs + postage + network overhead
= $0.12 + postage + network overhead

CONTENTS—UMST600

pages
1– 2 Identification & Abstract
3– 6 User Instructions
7– 9 I/O
11 Cost—Contents

DESCRIPTIVE TITLE	Chi Square from Raw Data
CALLING NAME	UMST620
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Unknown
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Decks and listing currently available
CONTACT	William Craig, EIN. Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This program computes and prints a frequency distribution of two variables based on intervals supplied to the program.

Optional outputs include:

1. A check of all the frequencies as calculated above to determine if all are equal to or above the given frequency supplied to the program. If cell frequencies are below this value, they are printed out and the problem terminated.
2. The chi-square statistic with its associated degrees of freedom.
3. Marginal sums of the calculated frequencies.
4. Expected frequencies based on the marginal sums and the total frequency.

Input data consist of observations to be grouped in a contingency table, commonly referred to as an r X c table with $r \geq 2$ and $c > 2$. Yates' continuity correction is applied in the case of 2 $\overline{X}$ 2 tables.

continued

Formulae

Let x_{ij} be the data element in the ith row and jth column.

Sums of columns: $n_{.j} = \sum_{i=1}^{r} x_{ij}$ where r is the number of rows.

Sums of rows: $n_{i.} = \sum_{j=1}^{c} x_{ij}$ where c is the number of columns.

Total: $N = \sum_{j=1}^{c} n_{.j}$

Chi square:

$$\text{Chi square} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(x_{ij} - \frac{n_{.j}n_{i.}}{N}\right)^2}{\frac{n_{.j}n_{i.}}{N}}$$

$$= N\left[\left(\sum_{i=1}^{r} \sum_{j=1}^{c} \frac{x_{ij}{}^2}{n_{i.}n_{.j}}\right) - 1\right]$$

The latter form is used in the program.

Degrees of freedom: d.f. = (r–1) (c–1)

Yates' continuity corrections:

Given the 2 X 2 table:

	a	b	a+b
	c	d	c+d
total	a+c	b+d	r

Use this table if (a+c)(a+b)/r = a

If (a+c)(a+b)/r > a, then use

a + 1/2	b – 1/2
c – 1/2	d + 1/2

instead of

a	b
c	d

continued

If $(a+c)(a+b)/r < a$, then use

$$\boxed{\begin{array}{ll} a - 1/2 & b + 1/2 \\ c + 1/2 & d - 1/2 \end{array}} \quad \text{instead of} \quad \boxed{\begin{array}{ll} a & b \\ c & d \end{array}}$$

Expected frequencies: $f_{ij} = n_{.j}n_{i.}/N$

USER INSTRUCTIONS

System Control Cards

System Control Cards will be prepared by Univ. of Minnesota personnel.

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8–10	NVAR	Number of variables to be read from Data Cards (does not include the optional counting variable). NVAR $\geq$ 2
11–15	NOBS	Number of observations or Data Cards to be read in. If blank or zero, an End of Data Card must be used and computer will count observations. NOBS $\geq$ 0
16–17	NCHI	Number of Chi-square Selection Cards following data. NCHI $\geq$ 1.
18	NAME	1: Name Cards provided 0: otherwise
19	MODE	Data Input Mode. 1: input of NVAR variables with no check for missing data. 2: input of NVAR variables with all blank and all non-numeric fields considered as missing data. *Warning:* Non-numeric characters are eliminated. Thus, 4.0 and 0.4 become 40 and 4 respectively. 3: input of NVAR variables after reading a missing data identification tag for each variable.
20	NOP1	1: output chi-square statistic 0: otherwise
21	NOP2	1: output the marginal sums 0: otherwise
22	NOP3	1: output expected frequencies 0: otherwise

continued

Columns	*Parameter*	*Contents*
23	XF	n: check all frequencies to be n or greater; printout of any which are not and suppression of chi-square calculation in those cases. $0 \leq n \leq 9$ 0: no check
24	NFC	Number of Format Cards. If zero or blank, NFC = 1 is assumed.
25–80	IDENT	Alphanumeric identification of the problem.

Format Card(s)

The Format Card must provide for reading one observation for each of NVAR variables using X, F, or E fields except in case of Input Mode 2 where A and X fields must be used. In addition, if NOBS is not given on the Problem Card, the first field read by the Format Card shall be an I field of at least two columns and an End of Data Card must be supplied. This field is used in counting observations. See End of Data Cards. If NFC = 0, NFC = 1 is assumed.

Name Card(s)

An alphabetic identification of 1 to 10 characters may be given to each variable by reading in Name Cards.

Columns	*Contents*
1–10	Name for variable 1
11–20	Name for variable 2
⋮	⋮
71–80	Name for variable 8

The format is repeated with 8 names to a card until each variable has been assigned a name.

Missing Data Identification Card(s) (Necessary for Input Mode 3)

Missing Data Identification Cards give a value for each variable, in accordance with the Format Card(s). If the value of a variable on a Data Card is the same as the value for that variable on the Missing Identification Card(s), the value of the variable on the data will be identified as missing. Each Data Card set is matched with the Missing Identification Card set to identify missing values.

continued

Input Data Cards

The Data Cards must be punched in accordance with the format given on the Format Card. One observation for each variable is read from a Data Card or Data Card set. Each observation starts on a new card or set of cards.

End of Data Card(s)

This card is required only if NOBS is not given on the Problem Card; i.e., NOBS is zero or blank. It must have –1 punched where it will be read by the first (i.e. integer) field of the Format Card. If the Format Card indicates that data is read from more than one card, the –1 card must be followed by enough blanks to form a data set. For example, if 4 cards are read in each data set, the –1 card should have 3 blank cards following it.

Chi-square Selection Cards

There must be as many Selection Cards as indicated by Cols. 16–17 of the Problem Card.

Columns	*Parameter*	*Contents*
1– 6		SELECT
7–10	NPROB	Identification number for this Selection Card.
11–14	NXX	Number of the X variable. $1 \le X \le NVAR$
15–18	NYY(1)	Number of the first Y variable if the solution mode (MODSEL) is 0 or 2. $0 \le NYY(1) \le NVAR$ (If NYY(1) = 0 or NYY(1) = NXX, no run is made.)
19–20	MODSEL	0: only one Y variable is to be selected with the given X variable and the given endpoints. 1: all variables other than NXX are run against NXX using the given endpoints. 2: up to ten additional variables may be run against NXX using the given endpoints.
21–24	XMIN	Minimum endpoint for X.
25–28	XMAX	Maximum endpoint for X. XMAX > XMIN The right-hand value of each interval is the endpoint, and the first and last intervals are open-ended. Thus,

continued

Columns	*Parameters*	*Contents*
25–28		if there are 2 intervals 1, 2, and 3, 4, then XMIN = 2, XMAX = 4, XDEL = 2. Values < 1 go into the first interval and values > 4 go into the last interval.
29–32	XDEL	Interval size for X; ($LX = \left[\frac{XMAX-XMIN}{XDEL} + 1\right]$ must be an integer ≥ 2, where LX is the number of intervals). XDEL > 0
33–36	YMIN	Minimum endpoint for Y.
37–40	YMAX	Maximum endpoint for Y. YMAX > YMIN The right-hand value of each interval is the endpoint, and the first and last intervals are open-ended.
41–44	YDEL	Interval size for Y; ($LY = \left[\frac{YMAX-YMIN}{YDEL} + 1\right]$ must be an integer ≥ 2, where LY is the number of intervals). YDEL > 0
	Note:	XMIN, XMAX, XDEL, YMIN, YMAX are read in F4.0 fields.
45–48	NYY(2)	The second Y variable to be run against the given X variable using the given endpoints if MODSEL is 2. $0 \leq NYY(2) \leq NVAR$ (If NYY(2) = 0 or NXX, no run is made.)
49–52	NYY(3)	The third variable to be run against the given X variable using the given endpoints if MODSEL is 2. $0 \leq NYY(3) \leq NVAR$ (If NYY(3) = 0 or NYY(3) = NXX, no run is made.)
⋮	⋮	
77–80	NYY(10)	The tenth Y variable to run against the given X variable using the given endpoints if MODSEL is 2. $0 \leq NYY(10) \leq NVAR$ (If NYY(10) = 0 or NYY(10) = NXX, no run is made.)

If MODSEL = 2, the number of NYY's to run against NXX is found by assuming there are 10 and ignoring those whose value of NYY(i) = 0 or NYY(i) = NXX.

continued

Use of the selection mode option (MODSEL) on the Selection Card should obviate the necessity of having a large number of Selection and Endpoint Cards.

Endpoint Cards

If X and/or Y are to be converted according to unequal intervals, give an array of endpoints on another card in the following manner; *ordered from smallest to largest.* The righthand value of each interval is the endpoint, and the first and last intervals are open-ended. If both X and Y are to be given this way, the X cards precede the Y cards. If either XDEL or YDEL or both are left blank on the Chi-square Selection Card an Endpoint Card is assumed for each blank value.

Columns	*Contents*
1– 6	ENDPTS
7– 8	LX or LY: number of endpoints following (2 $\le$ LX or LY).
9–12	Smallest endpoint.
13–16	Second endpoint.
⋮	⋮
77–80	18th endpoint.

Continue punching additional cards using four-column fields until all endpoints have been punched. Endpoints are read with F4.0 fields.
Caution: This program will allow an open frequency cell with no error printout. Make sure that the largest data element lies within the largest endpoint.

Finish Card

Columns	*Contents*
1– 6	FINISH

Used after last problem.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of the cards.

continued

```
SAMPLE INPUT
PROBLEM  4      401111 1UMST620 SAMPLE PROBLEM
(I2,4F4.0)
     2   1   9 107
     2   2   7 141
     2   2   4 135
     1   1   2 101
     2   2   4 110
     1   1   4 104
     2   1   9 111
     2   2   3 120
     1   1   4  97
     1   2   4  91
     1   1   3  87
     1   1   3  90
     2   2   8  93
     1   2   7  94
     1   1   7  90
-1
SELECT   1   1   2 0   1   2   1   1   2   1
SELECT   2   1   3 0   1   2   1
ENDPTS 2   4  10
SELECT6666   4   1 1               1   2   1
ENDPTS 2  94 200
SELECT6667   1   2 2   1   2   1   1   2   1   2   2   2
FINISH
```

SAMPLE OUTPUT

```
UMST620       CHI-SQUARE  FROM RAW DATA                        UMST620 SAMPLE PROBLEM

PROBLEM CARD

              4 VARIABLES
             -0 OBSERVATIONS
              4 SELECT CARDS (PROBLEMS)
           NAME CARDS  0 = NO
           MODE = 1
           CHI-SQUARE OUTPUT  1 = YES
           MARGINAL SUMS OUTPUT  1 = YES
           EXPECTED FREQUENCY OUTPUT  1 = YES
           FREQUENCY CHECK OUTPUT = -0
              1 FORMAT CARDS


           FORMAT = (I2,4F4.0)


           NUMBER OF OBSERVATIONS COUNTED =    15




SELECT CARD

           NPROB =    1   NXX =    1   NYY(1) =    2   MODSEL =  0
           XMIN =    1.0000   XMAX =    2.0000   XDEL =    1.0000   YMIN =    1.0000   YMAX =    2.0000   YDEL =    1.0000
           NYY(2 THRU 10) =    -0   -0   -0   -0   -0   -0   -0   -0   -0

VARIABLE       1 CONVERTED TO FREQUENCIES FOR VALUES FROM  1.000000E+00  TO 2.000000E+00 IN INTERVALS OF 1.000000E+00

VARIABLES NYY(I) CONVERTED TO FREQUENCIES FOR VALUES FROM 1.000000E+00  TO 2.000000E+00 IN INTERVALS OF 1.000000E+00

           THIS PROBLEM REQUIRES A FIELD LENGTH OF   027500


++++ PROBLEM    1   VARIABLE       1.             (X)   AND VARIABLE       2.             (Y)

OBSERVED FREQUENCIES

ROW  1        6       2

ROW  2        2       5

           CHI-SQUARE =       1.6370         DEGREES OF FREEDOM =    1


         ROW TOTALS         I =   2
                 8          7
```

continued

```
COLUMN TOTALS        J =   2
            8          7

GRAND TOTAL =          15


EXPECTED FREQUENCIES

ROW   1        4.267      3.733

ROW   2        3.733      3.267




SELECT CARD

          NPROB =    2   NXX =    1   NYY(1) =    3   MODSEL =  0
          XMIN =    1.0000   XMAX =    2.0000   XDEL =    1.0000   YMIN =   -0.0000   YMAX =   -0.0000   YDEL =   -0.0000
          NYY(2 THRU 10) =    -0   -0   -0   -0   -0   -0   -0   -0   -0

VARIABLE       1 CONVERTED TO FREQUENCIES FOR VALUES FROM  1.000000E+00  TO 2.000000E+00 IN INTERVALS OF 1.000000E+00

VARIABLES NYY(I) CONVERTED TO FREQUENCIES FOR   2 ENDPOINTS

 4.0000E+00  1.0000E+01

          THIS PROBLEM REQUIRES A FIFLD LENGTH OF    027500

++++ PROBLEM    2   VARIABLE       1.              (X)   AND VARIABLE       3.              (Y)

OBSERVED FREQUENCIES

ROW   1        6       3

ROW   2        2       4

          CHI-SQUARE =        .5469          DEGREES OF FREEDOM =    1
          .                   .
          .                   .
          .                   .

EXPECTED FREQUENCIES

ROW   1        4.800      4.200

ROW   2        3.200      2.800
```

continued

```
SELECT CARD
           NPROB = 6666   NXX =    4   NYY(1) =    1   MODSEL =  1
           XMIN =   -0.0000   XMAX =   -0.0000   XDEL =   -0.0000   YMIN =    1.0000   YMAX =    2.0000   YDEL =    1.0000
           NYY(2 THRU 10) =    -0   -0   -0   -0   -0   -0   -0   -0   -0

VARIABLE      4 CONVERTED TO FREQUENCIES FOR   2  ENDPOINTS
 9.4000E+01  2.0000E+02

VARIABLES NYY(I) CONVERTED TO FREQUENCIES FOR VALUES FROM 1.000000E+00  TO 2.000000E+00 IN INTERVALS OF 1.000000E+00

           THIS PROBLEM REQUIRES A FIELD LENGTH OF    027500

**** PROBLEM 6666   VARIABLE       4.               (X)   AND VARIABLE       1.              (Y)

OBSERVED FREQUENCIES
ROW  1          5        3
ROW  2          1        6

           CHI-SQUARE =        1.8862          DEGREES OF FREEDOM =    1


         ROW TOTALS            I =  2
                     8         7

         COLUMN TOTALS         J =  2
                     6         9

         GRAND TOTAL =           15


EXPECTED FREQUENCIES
ROW   1          3.200      4.800
ROW   2          2.800      4.200
```

continued

```
**** PROBLEM 6666   VARIABLE       4,              (X)   AND VARIABLE       2,              (Y)

OBSERVED FREQUENCIES

ROW  1        3       5

ROW  2        3       4

         CHI-SQUARE =         .1004          DEGREES OF FREEDOM =    1


        ROW TOTALS           I =   2

                  8          7

        COLUMN TOTALS        J =   2

                  6          9

        GRAND TOTAL =           15


EXPECTED FREQUENCIES

ROW   1        3.200       4.800

ROW   2        2.800       4.200




**** PROBLEM 6666   VARIABLE       4,              (X)   AND VARIABLE       3,              (Y)

OBSERVED FREQUENCIES

ROW  1        0       0

ROW  2        6       9

         CHI-SQUARE =        0.0000          DEGREES OF FREEDOM =    1


        ROW TOTALS           I =   2

                  0          15

        COLUMN TOTALS        J =   2

                  6          9

        GRAND TOTAL =           15


EXPECTED FREQUENCIES

ROW   1        0.000       0.000

ROW   2        6.000       9.000
```

continued

```
SELECT CARD

        NPROB = 6667   NXX =    1   NYY(1) =    2   MODSEL =  2
        XMIN =    1.0000   XMAX =    2.0000   XDEL =    1.0000   YMIN =    1.0000   YMAX =    2.0000   YDEL =    1.0000
        NYY(2 THRU 10) =     2    2    2   -0   -0   -0   -0   -0   -0

VARIABLE       1 CONVERTED TO FREQUENCIES FOR VALUES FROM  1.000000E+00  TO 2.000000E+00 IN INTERVALS OF 1.000000E+00

VARIABLES NYY(I) CONVERTED TO FREQUENCIES FOR VALUES FROM 1.000000E+00  TO 2.000000E+00 IN INTERVALS OF 1.000000E+00

        THIS PROBLEM REQUIRES A FIELD LENGTH OF   027500


**** PROBLEM 6667   VARIABLE       1.              (X)   AND VARIABLE       2.              (Y)

OBSERVED FREQUENCIES

ROW  1       6      2

ROW  2       2      5

        CHI-SQUARE =       1.6370          DEGREES OF FREEDOM =    1
      .              .
      .              .
      .              .
EXPECTED FREQUENCIES

ROW   1       4.267      3.733

ROW   2       3.733      3.267



**** PROBLEM 6667   VARIABLE       1.              (X)   AND VARIABLE       2.              (Y)

OBSERVED FREQUENCIES

ROW  1       6      2

ROW  2       2      5

        CHI-SQUARE =       1.6370          DEGREES OF FREEDOM =    1
      .              .
      .              .
      .              .
EXPECTED FREQUENCIES

ROW   1       4.267      3.733

ROW   2       3.733      3.267
```

continued

```
**** PROBLEM 6667   VARIABLE       1.              (X)   AND VARIABLE       2.              (Y)

OBSERVED FREQUENCIES
ROW  1         6       2
ROW  2         2       5

          CHI-SQUARE =        1.6370          DEGREES OF FREEDOM =    1


       ROW TOTALS          I =  2
                 8          7

       COLUMN TOTALS       J =  2
                 8          7

       GRAND TOTAL =          15


EXPECTED FREQUENCIES
ROW   1         4.267      3.733
ROW   2         3.733      3.267
```

```
**** PROBLEM 6667   VARIABLE       1.              (X)   AND VARIABLE       2.              (Y)

OBSERVED FREQUENCIES
ROW  1         6       2
ROW  2         2       5

          CHI-SQUARE =        1.6370          DEGREES OF FREEDOM =    1


       ROW TOTALS          I =  2
                 8          7

       COLUMN TOTALS       J =  2
                 8          7

       GRAND TOTAL =          15


EXPECTED FREQUENCIES
ROW   1         4.267      3.733
ROW   2         3.733      3.267
```

COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 0.371 seconds. At the current rate for the Univ. of Minnesota ($0.20/sec.), the computer time cost $0.07 plus a charge for output supplies.

Charge to user = computer costs + postage + network overhead

= $0.29 + postage + network overhead

CONTENTS—UMST620

pages	
1– 3	Identification & Abstract
5–10	User Instructions
11–17	I/O
19	Cost—Contents

DESCRIPTIVE TITLE	General Linear Hypothesis for Anova, Unequal Frequencies
CALLING NAME	UMST610
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Modified version of BMD05V from Health Sciences Comp. Fac., UCLA
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Lib., Minneapolis, Minn. 55455 Tel.: (612) 376-7232

FUNCTIONAL ABSTRACT

This program (a modification of UCLA's program BMD05V) analyzes the statistical significance of dependent variables for those experimental designs that can be formulated in terms of the General Linear Hypothesis model. It allows processing of the same design for as many dependent variables as field length will allow, where all variables are read in at once and the calculations are performed on the variables one at a time.

This program does *not* allow analysis of covariance. It does *not* allow transformation or transgeneration. The program can analyze unbalanced analysis of variance designs according to stated hypotheses.

Description of Model

Suppose we have n observations (values assumed by n random variables) $y_1, y_2, \dots, y_n$, which have the linear structure

$$y_\alpha = x_{1\alpha}\beta_1 + x_{2\alpha}\beta_2 + \dots + x_{\ell\alpha}\beta_\ell + e_\alpha,$$

$$\alpha = 1, 2, \dots, n$$

continued

where $\{x_{i\alpha}\}$ are known constants, $\{\beta_i\}$ are unknown variables, and $\{e_\alpha\}$ are uncorrelated random variables, with means 0 and variances σ^2.

The program is designed so that

$$x_{ij} = x_{i2}, x_{i2}, \ldots, x_{in}$$

is the design or analysis of variance type. The analysis of variance parameters are to be represented by the numbers 0, 1, –1 (or actually –9 to +9).

Using this program it is possible to process a great variety of analysis of variance designs; for example, full factorials, partial factorials, unbalanced designs, and experimental designs with missing values.

For a given set $y_\alpha, x_{1\alpha}, \ldots, x_{\ell\alpha}$ the program computes the least squares estimates,

$$\hat{\beta}_i, \quad i = 1, 2, \ldots, \ell,$$

the residual sum of squares

$$\sum_{\alpha=1}^{n} (y_\alpha - \sum_{i=1}^{\ell} x_{i\alpha}\hat{\beta}_i)^2,$$

and the degrees of freedom for this sum of squares. Hypotheses of the form,

$$H: \beta_{i_1} = \beta_{i_2} = \ldots = \beta_{i_k} = 0,$$

where $(i_1, i_2, \ldots, i_k)$ is any subset of $(1, 2, \ldots, \ell)$ may be stated. Least-squares estimates of the β's under these hypotheses are computed along with the residual sums of squares, and the degrees of freedom for these sums.

If we now make the assumption that the e_α are normally distributed, then the above information is sufficient to compute the usual F tests for the hypotheses stated, with the proper degrees of freedom.

USER INSTRUCTIONS

System Control Cards

System Control Cards will be prepared by Univ. of Minnesota personnel.

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
11–13	N	Number of variables to be read for processing under this design. $N \geq 1$
14–16	ND	Number of Design Cards (see Design Cards) (same as the number of cells). $ND \geq 2$
17–19	NP	Number of analysis of variance parameters (see Functional Abstract, Description of Model) (same as the number of cells, i.e., ND) $NP \geq 2$
20–22	NH	Number of Hypothesis Card Groups supplied to the program (see Hypothesis Card Groups) $NH \geq 0$
23–24	NFC	Number of variable Format Cards. If zero or blank, NFC = 1 is assumed. $NFC \geq 0$.
25–26	NA	01: test of accuracy of coefficients is desired. 00: test is not desired
27–80		Name (alphanumeric identification).

Format Card(s)

The Format Card provides for reading in one observation of the Data Cards for each variable with F or X–fields.

Design Cards

Columns	*Contents*
1– 4	R: the number of subjects or replicates in this cell.

continued

Columns	*Contents*
5– 6	00, 01, or –1 according to design specifications.
7– 8	00, 01, or –1 according to specifications.
⋮	⋮
79–80	00, 01, or –1 according to specifications. (Actually –9 to +9 are allowed as long as the Design Card columns add up to either 0 or ND.)

If there are more than 38 parameters, continue punching on following cards using 2 columns per parameter. Number of subjects is *not* repeated on following cards. (See Example and Card Preparation.)

Data Cards

On these cards are punched the values of the N dependent variables in accordance with the Format Card. The number of groups of N is R, as punched in Col. 1–4 on the Design Card.

Example: N = 2. Denote the two variables by x_i and y_i. Suppose the cell being considered has R = 3 replicates. The values of the variables could be punched on three cards, as X_1 Y_1 on the first, X_2 Y_2 on the second, X_3 Y_3 on the third, and read by a format such as (F3.0, F4.0), or could be punched on six cards, as

$$X_1;\ Y_1;\ X_2;\ Y_2;\ X_3;\ Y_3$$

and read by a format such as (F3.0/F4.0).

Hypothesis Card Groups

Columns	*Contents*
1	1 or 0 according to the hypothesis.
2	1 or 0 according to the hypothesis.
⋮	⋮
NP	1 or 0 according to the hypothesis.

The total number of ones and zeros punched must equal the number of analysis of variance parameters, NP, and the Hypothesis Card Group may be continued on more than one card with up to 80 values per card.

continued

The number of Hypothesis Card Groups supplied for the problem must agree with the number specified in Col. 20–22 of the Problem Card. (See Example and Card Preparation.)

Finish Card

Columns	*Contents*
1– 6	FINISH

This card is used after last problem.

Order of Cards

More than one problem can be processed consecutively. The following example illustrates the order of cards.

Program Control Cards for problem 1
   Problem Card for problem 1
   Format Card(s)
Design Cards and Data Cards for problem 1
   Design Card
   Data Cards corresponding to the Design Card
   Note: The combination of Design and Data Cards should be repeated for each cell in the problem.
Hypothesis Card Groups for problem 1
Program Control Cards for problem 2

   ⋮

Hypothesis Card Groups for problem 2

   ⋮

Program Control Cards for last problem

   ⋮

Design Cards and Data Cards for last problem

   ⋮

Hypothesis Card Groups for last problem

   ⋮

Finish Card

Length and Time Estimates

The problem length is a function only of the number of Design Cards and the number of Hypothesis Card Groups. Timing is dependent on the number of variables, N, on the number of Design Cards, ND, on the number of Hypothesis Card Groups, NH, and on the R–values (number of subjects in each cell) on each Design Card.

continued

Example and Card Preparation

Example

2 X 2 factorial with 3 replications and 2 missing values.

	B_1	B_2
A_1	y_{111} missing	$y_{112} = 6$
	$y_{211} = 5$	$y_{212} = 5$
	$y_{311} = 3$	$y_{312} = 7$
A_2	$y_{121} = 13$	$y_{122} = 12$
	$y_{221} = 14$	y_{222} missing
	$y_{321} = 15$	$y_{322} = 10$

The model is

$$E\,y_{rij} = m + a_i + b_j + d_{ij},$$

where a and b are variable effects, d is interaction effects, i = 1, 2, j = 1, 2, r = number of replicates;

$$\sum_{i=1}^{2} a_i = \sum_{j=1}^{2} b_j = \sum_{i=1}^{2} d_{ij} = \sum_{j=1}^{2} d_{ij} = 0.$$

This may be written as

$$Ey_{r11} = m + a_1 + b_1 + d_{11}\,, \quad r = 2, 3;$$

$$Ey_{r12} = m + a_1 + b_2 + d_{12}\,, \quad r = 1, 2, 3;$$

$$Ey_{r21} = m + a_2 + b_1 + d_{21}\,, \quad r = 1, 2, 3;$$

$$Ey_{r22} = m + a_2 + b_2 + d_{22}\,, \quad r = 1, 3;$$

and

$$a_2 = -a_1$$

$$b_2 = -b_1$$

continued

$d_{11} = -d_{12}$

$d_{21} = -d_{22}$

$d_{12} = -d_{22}$

Hence,

$$Ey_{r11} = 1\cdot m + 1\cdot a_1 + 1\cdot b_1 + 1\cdot d_{22}\,, \quad r = 2, 3$$

$$Ey_{r12} = 1\cdot m + 1\cdot a_1 - 1\cdot b_1 - 1\cdot d_{22}\,, \quad r = 1, 2, 3$$

$$Ey_{r21} = 1\cdot m - 1\cdot a_1 + 1\cdot b_1 - 1\cdot d_{22}\,, \quad r = 1, 2, 3$$

$$Ey_{r22} = 1\cdot m - 1\cdot a_1 - 1\cdot b_1 + 1\cdot d_{22}\,, \quad r = 1, 3$$

Preparation of Design Cards

The values that are to be entered on the Design Cards may be obtained from the last set of equations.

	Column											
Cell	1	2	3	4	5	6	7	8	9	10	11	12
11				2		1		1		1		1
12				3		1		1	−	1	−	1
21				3		1	−	1		1	−	1
22				2		1	−	1	−	1		1
				R		m		a_1		b_1		d_{22}

Note: Here, Col. 6 must sum to 4 and Cols. 8, 10, 12 must sum to zero.

Preparation of Data Cards

Following each Design Card(s) there must be one or more Data Cards containing the values of the dependent variables for every replicate specified in the first four columns of this Design Card. The Data Cards in this case will have the format of (F3.0) Let us assume that least square estimates of the a_i, b_j, d_{ij} are desired. Then the cards are arranged as follows:

continued

```
                      Column
  1  2  3  4  5  6  7  8  9 10 11 12 13 14 ...
           2     1     1     1     1              Design Card 1
        5
        3
           3     1     1  -  1  -  1              Design Card 2
        6
        5
        7
           3     1  -  1     1  -  1              Design Card 3
     1  3
     1  4
     1  5
           2     1  -  1  -  1     1              Design Card 4
     1  2
     1  0
```

Preparation of Hypothesis Card Groups

Each Hypothesis Card is used to define a hypothesis on the β_i, $i = 1, 2, \ldots, \ell$. In our example the hypothesis of interest are:

1. $a_1 = a_2 = 0$ or equivalently, $a_1 = 0$
2. $b_1 = b_2 = 0$ or equivalently, $b_1 = 0$
3. $d_{11} = d_{12} = d_{21} = d_{22} = 0$ or equivalently, $d_{22} = 0$

The Hypothesis Card Groups corresponding to the hypotheses 1–3 above are:

Hypothesis	*Columns*				*Effect Tested*
	1	2	3	4	
1	1	0	1	1	Column
2	1	1	0	1	Row
3	1	1	1	0	Interaction
	m	a	b	d	

In general, if the hypothesis

$$H: \beta_{i_1} = \beta_{i_2} = \ldots = \beta_{i_k} = 0$$

continued

is of interest, simply enter a Hypothesis Card Group with zeros appearing in columns i_1, i_2, ..., i_k and ones in the remaining ℓ-k columns.

Information Provided by the Program

The program lists the various hypotheses to be analyzed, using the zero-one code described in the Preparation of Hypothesis Card Groups. These hypotheses are listed in a table titled HYPOTHESES. The first two rows of the table are:

1. 00000...0
2. 11111...1

The last row is

h + 3 10000...0,

where h is the number of Hypothesis Card Groups. These three hypotheses are automatically included and need not be prepared by the user.

Following this table there appears the table titled ESTIMATE OF COEFFICIENTS, RESIDUAL SUMS OF SQUARES, AND F-TESTS.

Missing Data Note

In cases where data are missing (y_{111}, y_{222} in this example), note that Data Cards were *not* prepared for the missing data. Also note that the value of R (Col. 4 on the Design Cards) is used to indicate the amount of data present for each design.

SAMPLE INPUT

```
PROBLEM     1  4  4  3 1  TEST PROBLEM
(F3.0)
   2 1 1 1 1
  5
  3
   3 1 1-1-1
  6
  5
  7
   3 1-1 1-1
 13
 14
 15
   2 1-1-1 1
 12
 10
1011
1101
1110
FINISH
```

SAMPLE OUTPUT

```
UMST610     ANOVA--GENERAL LINEAR HYPOTHESIS                          TEST PROBLEM

        1 VARIABLES
        4 DESIGN CARDS
        4 PARAMETERS
        3 HYPOTHESIS CARDS
        1 FORMAT CARDS
     TEST OF ACCURACY OF COEFFICIENTS    -0 = NO

          FORMAT = (F3.0)


DESIGN PATTERNS

   1    1 1 1 1
   2    1 1-1-1
   3    1-1 1-1
   4    1-1-1 1

          THIS PROBLEM REQUIRES A FIELD LENGTH OF ONLY 033500B


..................VARIABLE NO.  1...........................


          NO. OF        MEAN         STD. DEV.
          CASES          Y              Y
   1        2          4.00000        1.41421
   2        3          6.00000        1.00000
   3        3         14.00000        1.00000
   4        2         11.00000        1.41421


HYPOTHESES AND SUMS OF SQUARES EXPLAINED BY HYPOTHESES

   1      0000          0.00000
   2      1111        970.00000
   3      1011        835.00000
   4      1101        969.40000
   5      1110        955.00000
   6      1000        810.00000


ESTIMATES OF COEFFICIENTS, RESIDUAL SUMS OF SQUARES, AND F TESTS

HYPOTHESIS        1             2             3             4             5             6
VARIABLE
    1          0.00000       8.75000       8.75000       8.75000       9.00000       9.00000
    2          0.00000      -3.75000       0.00000      -3.80000      -3.75000       0.00000
    3          0.00000        .25000       1.00000       0.00000        .25000       0.00000
    4          0.00000      -1.25000      -1.25000      -1.25000       0.00000       0.00000

RESIDUAL
SUM SQS.     978.00000       8.00000     143.00000       8.60000      23.00000     168.00000

DEGREES OF
FREEDOM OF      10             6             7             7             7             9
RESIDUALS

F TESTS                                  101.25000        .45000      11.25000      40.00000

DEGREES OF
FREEDOM OF                                 1    6        1    6        1    6        3    6
F TESTS
```

COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 0.164 seconds. At the current rate for the Univ. of Minnesota ($0.20/sec.), the computer time cost $0.03 plus a charge for output supplies.

Charge to user = computer costs + postage + network overhead
= $0.07 + postage + network overhead

CONTENTS—UMST610

pages
1– 2 Identification & Abstract
3– 9 User Instructions
11–12 I/O
13 Cost—Contents

DESCRIPTIVE TITLE	Single and Simultaneous Equation (TSLS, LISE) Regression Package
CALLING NAME	UMST630
INSTALLATION NAME	University of Minnesota University Computer Center
AUTHOR(S) AND AFFILIATION(S)	Morris Norman University of California, Santa Barbara James Weatherby University of Texas
LANGUAGE	CDC Fortran IV
COMPUTER	CDC6600 (Scope 3.1.6)
PROGRAM AVAILABILITY	Deck and listing currently available
CONTACT	William Craig, EIN Tech. Rep., Center for Urban and Regional Affairs, Univ. of Minn., 311 Walter Library, Minneapolis, Minn. 55455 Tel.: (612) 373-7833

FUNCTIONAL ABSTRACT

This is a multipurpose program designed to estimate the coefficients of a multiple regression model or a simultaneous equation model. Three estimation techniques are available to the user: ordinary least squares (OLS)[1,2]; two stage least squares (TSLS)[3,4]; and limited information maximum likelihood (LISE)[5,6]. The program has variable input format and a transformation routine with nine options including a provision to lag variables one period. For the computational methods used see Ref. 7.

The standard printout for each equation estimated is as follows:

a. Augmented moment matrix (raw cross-product matrix)
b. Variance-covariance matrix of the coefficients
c. Regression coefficients and their asymptotic standard errors
d. Variance and standard deviation of the equation
e. R^2 adjusted for degrees of freedom
f. R^2 not adjusted for degrees of freedom
g. t test for each regression coefficient; H_0: $B = 0$, H_1: $B \neq 0$
h. F test; H_0: $(B_1,\dots,B_n) = 0$ H_1: $(B_1,\dots,B_n) \neq 0$.

continued

The optional printout for each equation estimated is as follows:

a. Covariance and correlation matrix for all of the variables
b. Graph of actual and predicted value of the dependent variable versus the corresponding observation
c. Durbin-Watson statistic
d. Scatter diagrams of values of the dependent variable plotted against the values of selected independent variable(s)
e. Elasticities of each independent variable with respect to the dependent variable (taken about the means)
f. Standard partial regression coefficient for each independent variable
g. If the transformation routine is used, a table describing the indicated transformation is printed. The number of transformations is restricted to forty.

Formulae of Selected Statistics

Moment Matrix of all the variables (optional)

$$= \frac{\sum_{k=1}^{n} Z_{ik}Z_{jk} - \frac{n\ \overline{Z}_i\ \overline{Z}_j}{(n-1)}}{n}$$

Correlation Matrix of all the variables (optional)

$$= \frac{\text{Moment Matrix}}{(\sigma_i \text{ X } \sigma_j)}$$

Variance–Covariance Matrix of Coefficient[8]

$$S_{cb,cb} = S^2 \frac{Y_1'\ X(X'X)^{-1}X'Y_1 \quad Y_1'X_1}{X_1'Y_1 \qquad\qquad X_1'X_1}$$

This is the analogue of E(β – Eβ)(β – Eβ). The standard errors given are the square roots of the diagonal elements of the matrix (these are given in parentheses).

Unadjusted R^2

$$R^2 = 1 - \frac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{\sum_{t=1}^{n} (y_t - \overline{y})^2}$$

continued

Adjusted R^2

$$\tilde{R}^2 = 1 - \frac{\dfrac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n - m - 1}}{\dfrac{\sum_{t=1}^{n} (y_t - \overline{y})^2}{n - 1}}$$

$$\beta_j = \frac{\hat{\beta}_j \ SE_j}{\sqrt{\dfrac{\sum_{t=1}^{n} (y_t - \hat{y}_t)^2}{n - m - 1}}}$$

Elasticity of y_t with respect to X_j

$$X_j = \frac{\overline{X}_j}{\overline{y}} \hat{\beta}_j$$

Symbols:

Z_j = the dependent or independent variable

y_t = dependent variable

$\hat{y}_t$ = estimated value of y_t

$\overline{y}$ = mean of y_t

n = number of observations

m = number of independent variables

$\hat{\beta}_j$ = estimate of β_j, the regression coefficient of the jth independent variable

SE_j = the standard error of $\hat{\beta}_j$

$\overline{X}_j$ = the mean of the jth independent variable

continued

REFERENCES

1. Goldberger, A.S., *Econometric Theory* (New York: John Wiley & Sons, Inc., 1964), Ch. 4.

2. Johnston, J., *Econometric Methods* (New York: McGraw Hill, Inc., 1963), Ch. 4.

3. Goldberger, pp. 329–336.

4. Johnston, pp. 258–260.

5. Goldberger, pp. 338–345.

6. Johnston, pp. 254–258.

7. Christ, C.F., *Econometric Models and Methods* (New York: John Wiley & Sons, Inc., 1966), Ch. 9, Sec. 9–10.

8. Goldberger, p. 333.

USER INSTRUCTIONS

System Control Card

System Control Cards will be prepared by Univ. of Minnesota personnel.

Program Control Cards

Problem Card

Columns	*Parameter*	*Contents*
1– 7		PROBLEM
8–11	ND	Number of observations
12–13	MN	Number of variables read in
14–15	NX	Number of exogenous variables (use only if running TSLS or LISE). Do not count the constant term. This is the final number of exogenous variables following transgeneration.
16–17	N1	01: output moment matrix and correlation matrix for all of the variables (those read off cards plus those generated by the transformation routine). blank: this option is not desired.
18–19	N2	Number of NEW variables ADDED by the transformations ($1 \le NZ \le 40$). blank: no transformations are necessary.
20–21	N3	01: running TSLS or LISE and the variables were not read in the correct order, i.e., *exogenous first, then endogenous*. blank: the variables were read in the correct order or running only OLS.
22–23	NFC	Number of Variable Format Cards. If zero or blank, NFC = 1 is assumed. $NFC \ge 0$
24–25	NEAM	01: output the elasticities of the dependent variable for each independent variable. blank: this option is not desired.

continued

Columns	*Parameter*	*Contents*
26–27	NSPRC	01: output the standard partial regression coefficients. blank: this option is not desired.
28–78	IDENT	Alphanumeric description of the problem.

Format Cards

Punch the format of the data to be read from cards, starting in Col. 1 of the first card and continuing on additional cards, until finished. The format must be preceded by a left parenthesis and followed by a right parenthesis. The number of cards the format occupies is the parameter in Cols. 22–23 of the Problem Card.

Input Data Cards

Punch the input data according to the format specified on the Format Cards.

Transformation Cards

Options Available	Transformation Code
Variable + variable	0001
Variable – variable	0002
Variable X variable	0003
Variable / variable (variable $\neq$ 0)	0004
Variable + constant	0005
Variable X constant	0006
Natural log of variable (variable > 0)	0007
Log_{10} of variable (variable > 0)	0008
Lagged variable	0009

Use of the Transformation Routine

The use of transformations requires specifying to the routine three pieces of information: the transformation code, the variable number of the variable to be created, and the variable number(s) and/or the constant being used to form the new variable. Specifically, transformation codes 0001–0004 require specification of two variable numbers, codes 0005 and 0006 require specification of one variable number and a constant, and codes 0007–0009 require specification of one variable number.

Each Transformation Card provides for all of the above parameters. However, all of the information need not be punched on the

continued

Transformation Card to specify the information for a particular transformation. The only information needed for every TRANS Card is NOPTN and NVARN, the remaining fields can be specified as needed.

Columns	*Parameter*	*Contents*
1– 5		TRANS
6– 8		blank
9–12	NOPTN	Transformation code
13–16	NVARN	Resulting variable number
17–20	NV1	First variable used in transformation
21–24	NV2	Second variable used in transformation
25–34	CONST	Constant used in transformation

Examples

Assume that 10 variables and 20 observations have been read and that the following transformations are needed (let V stand for variable).

1. Lag V1 one period
2. Lag V7 one period
3. Subtract V7 from V9
4. Add V3 and V4
5. Deflate V6 by 1000
6. Take the natural log of V8.

Since six transformations are called for, the parameter in Cols. 18–19 if the Problem Card should equal 6.

Assume that all of the variables used to form the new variables are needed in the analysis except V6. In the case of V6, thus, it is not necessary to create a new variable. Therefore simply replace V6 by V6/1000. Since the rest of the original variables are needed in the analysis, new variables must be created as the necessary transformations are performed. Symbolically the transformation will look as follows:

$V11 = V1_{-1}$

$V12 = V7_{-1}$

$V13 = V7-V9$
$V14 = V3+V4$
$V6 = V6/1000.$
$V8 = \log_e(V8)$

continued

740

and the TRANS Cards will be:

	NOPTN	NVARN	NV1	NV2	CONST[α]
TRANS	0009	0011	0001	blank	blank
TRANS	0009	0012	0007	blank	blank
TRANS	0002	0013	0007	0009	blank
TRANS	0001	0014	0003	0004	blank
TRANS	0004	0006	0006	blank	1000.
TRANS	0008	0015	0008	blank	blank

Note: V13 could not be used in an equation with both V7 and V9 since the cross-product matrix would then be singular. The same is true of V14 with V3 and V4.

[α] The input format for the CONST. is F10.0 thus the decimal *must be punched.*

As will be recalled, when lagging variables one observation is lost. The program takes this into account. Thus after the transformations have been completed there will be 15 variables (the original nine, one replaced, five new) and 19 observations. With respect to the lagged variable, the last card is the newest reading. This transformation routine only allows one period lags, more than this must be placed as a separate variable on the Data Cards.

The program will print the data read from cards, then a table indicating the desired transformations, and finally the data after the transformations.

The following is included as an example of how the user may make more complicated transformations. Ten variables and 40 observations are read in; and the following transformation is desired:

$$V11 = [(V2 - V6 + V7)/V4]_{-1}$$

This new variable requires 4 TRANS Cards as follows:

	NOPTN	NVARN	NV1	NV2	CONST
TRANS	0002	0011	0002	0006	blank
TRANS	0001	0011	0011	0007	blank
TRANS	0004	0011	0011	0004	blank
TRANS	0009	0011	0011	blank	blank

It will be noted that V11 is used as a dummy storage location for the first three transformations. The fourth transformation creates the desired variable and stores the result in V11 which can now be used in the analysis.

continued

Label Card(s)

Columns	*Contents*
1– 5	LABEL
6–10	First variable name
11–15	Second variable name
⋮	⋮
76–80	15th variable name

If more than 15 variables are used, the names are continued on the next card as follows.

Columns	*Contents*
1– 5	LABEL
6–10	16th variable name
⋮	⋮

Location Card

Necessary only if Cols. 20–21 of Problem Card = 01.

Columns	*Contents*
1– 6	LOCATN
7– 8	blank
9–10	In fields of two, punch the variable number of each exogenous variable, then the location of the constant term (the total number of variables plus 1), and then the variable numbers of each of the endogenous variables.
11–12	
⋮	

Analysis Card(s)

Columns	*Contents*
1– 6	ANALYS
7	Estimating option 0: OLS 1: TSLS 2: TSLS and LISE
8	1: this is last ANALYS Card for this problem blank or zero: otherwise

continued

Columns	*Contents*
9–10	Number of endogenous variables minus one in equation (i.e., number of $\hat{y}$'s) if running TLSL or LISE.
11–12	Number of variables in this equation not including constant.
13– x	In fields of 2, punch the number of the field in which the variables in this equation appeared in the input data matrix. For OLS, the order is the independent variables' location, then the dependent variable's location. For TSLS and LISE, the explanatory endogenous variables' locations, then the exogenous variables' location, i.e., in terms of TSLS notation *the order is $\hat{y}$'s, the exogenus, then the dependent y.*
x+1,x+2	This field is the next field of 2 after the location numbers. 01: graph of actual versus predicted and the Durbin–Watson d statistic are output 02: Durbin-Watson d statistic only is output 03: graph of actual versus predicted and the Durbin–Watson d statistic plus scatter diagrams of the dependent variable against up to 10 independent variables are output. In this case, in Cols. x+5, x+6 punch the number of variables and then their locations in fields of two.
x+3,x+4	01: suppress constant term blank: otherwise
x+5,x+6	01: scatter diagram of dependent with first independent is output. However, if Cols. x+1, x+2 = 03, punch the number of variables to be scattered against the dependent variable, then their locations (still in fields of two).

After specifying the entire analysis, there may not be enough room on one ANALYS Card. If this is the case, simply finish specifying the analysis on a second ANALYS Card as follows.

Columns	*Contents*
1– 6	ANALYS
7– x	Fill in the remaining information in two-column fields.

continued

Finish Card

Columns	*Contents*
1– 6	FINISH

Used after last problem.

Order of Cards

More than one problem may be processed consecutively. The following example illustrates the order of cards.

Program Control Cards
- Problem Card
- Format Card(s)
- Data Cards
- Transformation Card(s) (if necessary)
- Label Card(s)
- Location Card (if necessary)
- Analysis Card(s)

Repeat for additional problems

Finish Card

SAMPLE INPUT

```
PROBLEM  11 4 2 1 1 1 1 1 1  THIS IS JOHNSON PR.PAGF 269. TEST OF PROGRAM
(5X,4(F5.0,2X))
1348 13895  10706   1024   2165      1
1949 14377  10940   1078   2359      1
1950 14843  11250   1123   2470      1
1951 15307  11089   1052   3166      1
1952 15360  11023    980   3357      1
1953 15951  11474   1073   3404      1
1954 16680  12023   1281   3376      1
1955 17237  12443   1474   3320      1
1956 17547  12548   1591   3408      1
1957 17788  12802   1668   3318      1
1958 17699  13096   1709   2894      1
TRANS      9   5   3
LABELGDOPRCONEXGFCF EXOG LAG1
LOCATN   4 5 6 1 2 3
ANALYS2  1 2 1 2 3 0 1 1
ANALYS2  1 3 1 5 3 3 0 2 1 5
FINISH
```

SAMPLE OUTPUT

```
     UMST630 -- ECONOMETRICS PROGRAM   THIS IS JOHNSON PR.PAGE 269, TEST OF PROGRAM
NUMBER OF OBSERVATIONS=   11NUMBER OF VARIABLES=  4NUMBER OF FORMAT CARDS=  1
NUMBER OF EXOGENOUS VARIABLES=  2 NUMBER OF TRANSFORMATIONS=  1
        OPTIONS
MOMENT MATRIX AND CORRELATION MATRIX FOR ALL THE VARIABLES YES.
VARIABLES NOT IN ASSUMED ORDER                    YES.
ELASTICITIES YES.
PARTIAL REGRESSION COEFFICIENTS YES.
   FORMAT=   (5X,4(F5.0,2X))

         THIS PROBLEM REQUIRES A FIELD LENGTH OF 041000B

          1            2            3            4
  1    13895.00     10706.00     1024.00      2165.00
  2    14377.00     10940.00     1078.00      2359.00
  3    14843.00     11250.00     1123.00      2470.00
  4    15307.00     11089.00     1052.00      3166.00
  5    15360.00     11023.00      980.00      3357.00
  6    15951.00     11474.00     1073.00      3404.00
  7    16680.00     12023.00     1281.00      3376.00
  8    17237.00     12443.00     1474.00      3320.00
  9    17547.00     12548.00     1591.00      3408.00
 10    17788.00     12802.00     1668.00      3318.00
 11    17699.00     13096.00     1709.00      2894.00




                                    TRANSFORMATION TABLE

            OPTION                 RESULTING VARIABLE                FORMED BY
                                          NUMBER              FIRST VARIABLE    SECOND VARIABLE        CONSTANT


       LAGGED VARIABLE                      5                        3                -0               -0.000000


A ZERO UNDER THE HEADING FORMED BY MEANS THAT THIS ITEM IS NOT USED IN THE OPTION IN QUESTION
```

continued

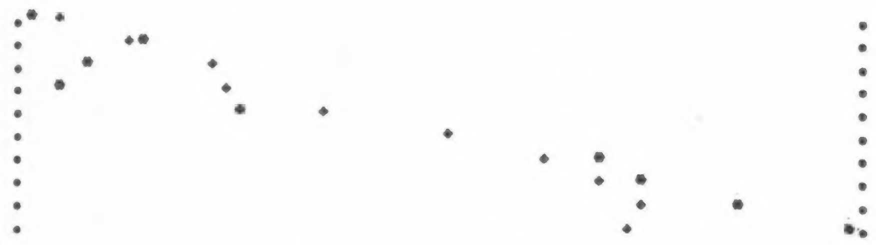
```
            1           2           3           4           5
  1    14377.00    10940.00     1078.00     2359.00     1024.00
  2    14843.00    11250.00     1123.00     2470.00     1078.00
  3    15307.00    11089.00     1052.00     3166.00     1123.00
  4    15360.00    11023.00      980.00     3357.00     1052.00
  5    15951.00    11474.00     1073.00     3404.00      980.00
  6    16680.00    12023.00     1281.00     3376.00     1073.00
  7    17237.00    12443.00     1474.00     3320.00     1281.00
  8    17547.00    12548.00     1591.00     3408.00     1474.00
  9    17788.00    12802.00     1668.00     3318.00     1591.00
 10    17699.00    13096.00     1709.00     2894.00     1668.00
LOCATION CARD
   4   5   6   1   2
```

```
 MOMENT MATRIX

         GDOPR         CONEX         GFCF          EXOG          LAG1
GDOPR  1.61739E+06
CONEX  9.93766E+05  6.57399E+05
GFCF   3.27796E+05  2.23262E+05  7.93317E+04
EXOG   2.95827E+05  1.13104E+05  2.52024E+04  1.57520E+05
LAG1   2.69075E+05  1.84809E+05  6.75117E+04  1.67546E+04  6.44367E+04
CONS   1.62789E+04  1.18688E+04  1.30290E+03  3.10720E+03  1.23440E+03
```

```
CORRELATION MATRIX

       GDOPR CONEX GFCF  EXOG
CONEX   .96
GFCF    .92   .98
EXOG    .59   .35   .23
LAG1    .83   .90   .94   .17

THE EXOGENOUS VARIABLES FOR THIS SYSTEM ARE
EXOG  LAG1  CONS


THE ENDOGENOUS VARIABLES ARE
GDOPR CONEX GFCF
```

continued

```
EQUATION NUMBER          1    NUMBER OF OBSERVATIONS   10
 TOTAL NUMBER OF VARIABLES    5       THIS IS JOHNSON PR.PAGE 269, TEST OF PROGRAM
ESTIMATING OPTION  2
NUMBER OF Y HATS IN THIS EQUATION IS  1
NUMBER OF VARIABLES IN THIS EQUATION IS  2
GRAPHING AND THE DURBIN WATSON OPTION IS  3
SUPPRESS CONSTANT TERM  0
VARIABLES USED IN THE SCATTER DIAGRAM
 1

 LIMITED-INFORMATION SINGLE EQUATION
 NUMBER OF ITERATION    9                 .26176             .26176


             1           2           3
   1  2.66314E+09  1.62789E+05  1.93992E+09
   2  1.62789E+05  1.00000E+01  1.18688E+05
   3  1.94105E+09  1.18688E+05  1.41460E+09
    VARIANCE-COVARIANCE MATRIX OF COEFFICIENT

  7.41708E-03
 -1.20742E+02  1.97527E+06


   DEPENDENT VARIABLE      CONEX


    A( GOOPR  )=       .45789280     T=  5.317     BETA COEF. =       .00012646     ELASTICITY(ABOUT THE MEAN)=      .62803241
               (       .08612246)


    A( CONS   )=  4414.80891838      T=  3.141
               ( 1405.44245474)


   VARIANCE= 9,72384284E+04

   STAND DEV=3.11830769E+02

   R-SQUARED=     .8521          (ADJUSTED FOR D. OF FREEDOM)

   F-TEST(  1,  8)=   52.8462

   R-SQUARED=     .8685          (NOT ADJUSTED FOR DEGREES OF FREEDOM)



    ACTUALS       PREDICTED       RESIDUALS      0/0 ERROR                              RANGE 10940.00  TO 13096.00

   10940.000      10997.934        -57.934          -.530    1
   11250.000      11211.312         38.688           .344    2
   11089.000      11423.774       -334.774         -3.019    3
   11023.000      11448.042       -425.042         -3.856    4
   11474.000      11718.657       -244.657         -2.132    5
   12023.000      12052.461        -29.461          -.245    6
   12443.000      12307.507        135.493          1.089    7
   12548.000      12449.454         98.546           .785    8
   12802.000      12559.806        242.194          1.892    9
   13096.000      12519.054        576.946          4.406   10
DURBIN-WATSON D STATISTIC=     .51044
```

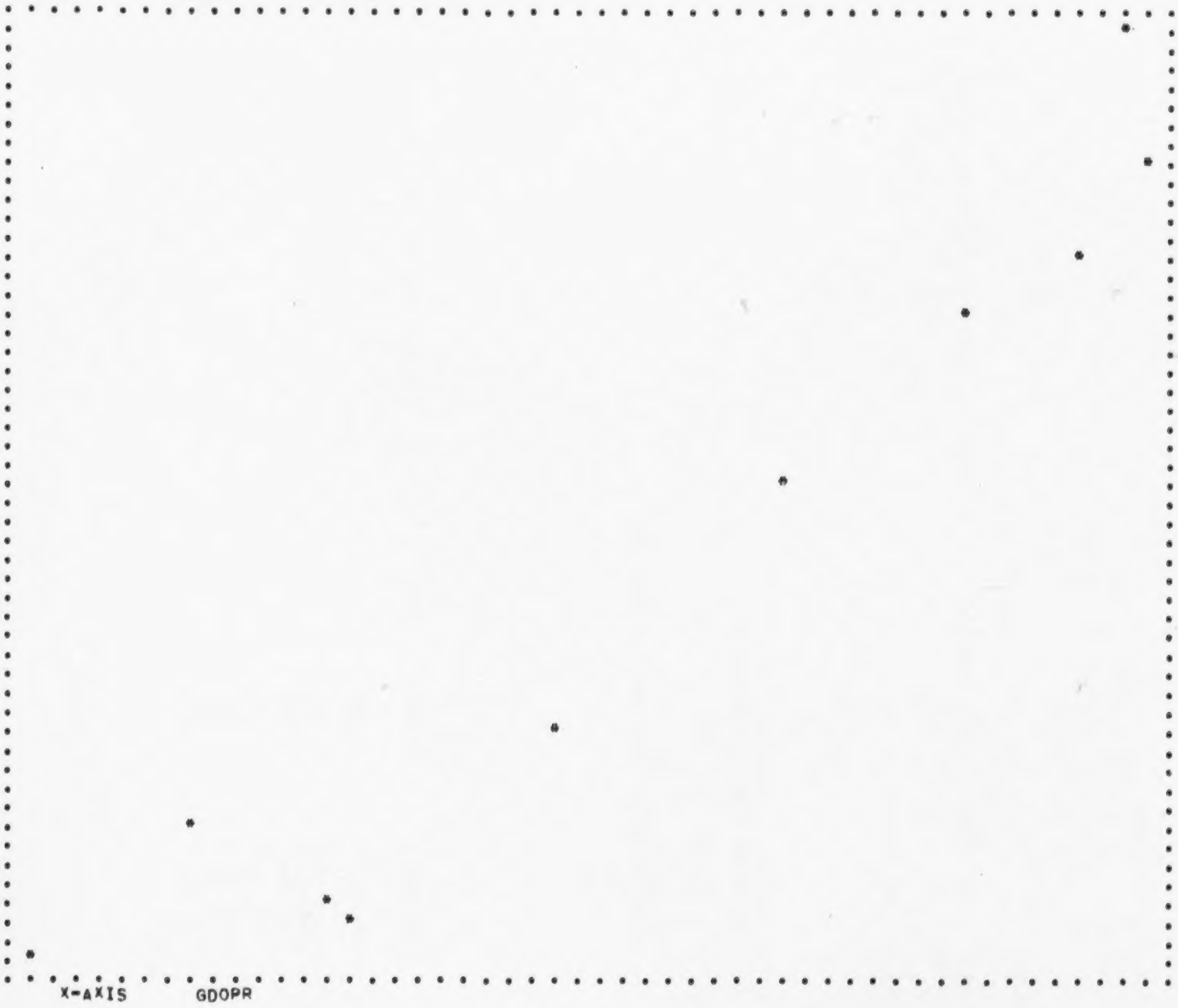


continued

```
EQUATION NUMBER          1     NUMBER OF OBSERVATIONS   10
 TOTAL NUMBER OF VARIABLES    5        THIS IS JOHNSON PR.PAGE 269. TEST OF PROGRAM

 TWO STAGE LEAST SQUARES

     VARIANCE-COVARIANCE MATRIX OF COFFFICIENT

  4.06531E-03
 -6.61788E+01  1.08265E+06

    DEPENDENT VARIABLE     CONEX

     A( GDOPR  )=       .59563978      T=  9.342     HETA COEF. =      .00016451     FLASTICITY(ABOUT THE MEAN)=      .81696216
               (       .06375979)

     A( CONS   )=  2172.43954157      T=  2.088
               ( 1040.50351171)

    VARIANCE= 5.32965252E+04

    STAND DEV=2.30860402E+02

    R-SQUARED=     .9189         (ADJUSTED FOR D. OF FREEDOM)

    F-TEST(  1,  8)=  103.0127

    R-SQUARED=     .9279         (NOT ADJUSTED FOR DEGREES OF FREEDOM)


    ACTUALS       PREDICTED       RFSIDUALS      0/0 ERROR                        RANGE 10735.95   TO 13096.00

   10940.000      10735.953        204.047         1.865    1
   11250.000      11013.521        236.479         2.102    2
   11089.000      11289.898       -200.898        -1.812    3
   11023.000      11321.467       -298.467        -2.708    4
   11474.000      11673.490       -199.490        -1.739    5
   12023.000      12107.711        -84.711         -.705    6
   12443.000      12439.482          3.518          .028    7
   12548.000      12624.131        -76.131         -.607    8
   12802.000      12767.680         34.320          .268    9
   13096.000      12714.668        381.332         2.912   10
  DURBIN-WATSON D STATISTIC=     .87150
```

continued

```
SCATTER DIAGRAM WITH VARYING SCALE
Y-AXIS  CONEX   RANGE       10940.000      13096.000
X-AXIS  GDOPR   RANGE       14377.000      17788.000
CODE OF SYMBOL   *=1   +=2   $=OVER 4
```



continued

```
EQUATION NUMBER          2    NUMHER OF OBSERVATIONS   10
 TOTAL NUMBER OF VARIABLES    5       THIS IS JOHNSON PR.PAGE 269. TEST OF PROGRAM
ESTIMATING OPTION  2
NUMBER OF Y HATS IN THIS EQUATION IS  1
NUMBER OF VARIABLES IN THIS EQUATION IS  3
GRAPHING AND THE DURBIN WATSON OPTION IS  3
SUPPRESS CONSTANT TERM  0
VARIABLES USED IN THE SCATTER DIAGRAM
 1  5

 LIMITED-INFORMATION SINGLE EQUATION
 NUMBER OF ITERATION    2       720732880605.27344  720732880604.96484


              1            2            3            4
    1  2.66314E+09  2.03368E+08  1.62789E+05  2.14737E+08
    2  2.03368E+08  1.58174E+07  1.23440E+04  1.66906E+07
    3  1.62789E+05  1.23440E+04  1.00000E+01  1.30290E+04
    4  2.15048E+08  1.66906E+07  1.30290E+04  1.76895E+07
     VARIANCE-COVARIANCE MATRIX OF COFFFICIENT

  2.59799E-03
 -1.08487E-02  5.87310E-02
 -2.89008E+01  1.04108E+02  3.42742F+05


    DEPENDENT VARIABLE     GFCF


     A( GDOPR  )=      .03386220     T=   .664     BETA COEF. =      .00001956    ELASTICITY(ABOUT THE MEAN)=      .42308653
               (       .05097052)


     A( LAG1   )=      .90631927     T=  3.740     BETA COEF. =      .00248890    ELASTICITY(ABOUT THE MEAN)=      .85866952
               (       .24234480)


     A( CONS   )=   -367.09994743    T=  -.627
               (     585.44173162)

    VARIANCE= 7.78781074E+03

    STAND DEV=8.82485736E+01

    R-SQUARED=     .9018          (ADJUSTED FOR D. OF FREEDOM)

    F-TEST(  2,  7)=   42.3399


    R-SQUARED=     .9236          (NOT ADJUSTED FOR DEGREES OF FREEDOM)

    ACTUALS       PREDICTED      RESIDUALS      0/0 ERROR                              RANGE   980.00   TO   1743.97


   1078.000       1047.808        30.192         2.801    1
   1123.000       1112.529        10.471          .432    2
   1052.000       1169.025      -117.025       -11.124    3
    980.000       1106.471      -126.471       -12.905    4
   1073.000       1061.229        11.771         1.097    5
   1281.000       1170.202       110.798         8.649    6
   1474.000       1377.578        96.422         6.542    7
   1591.000       1562.995        28.005         1.760    8
   1668.000       1677.195        -9.195         -.551    9
   1709.000       1743.968       -34.968        -2.046   10

 DURBIN-WATSON D STATISTIC=     .96463
```

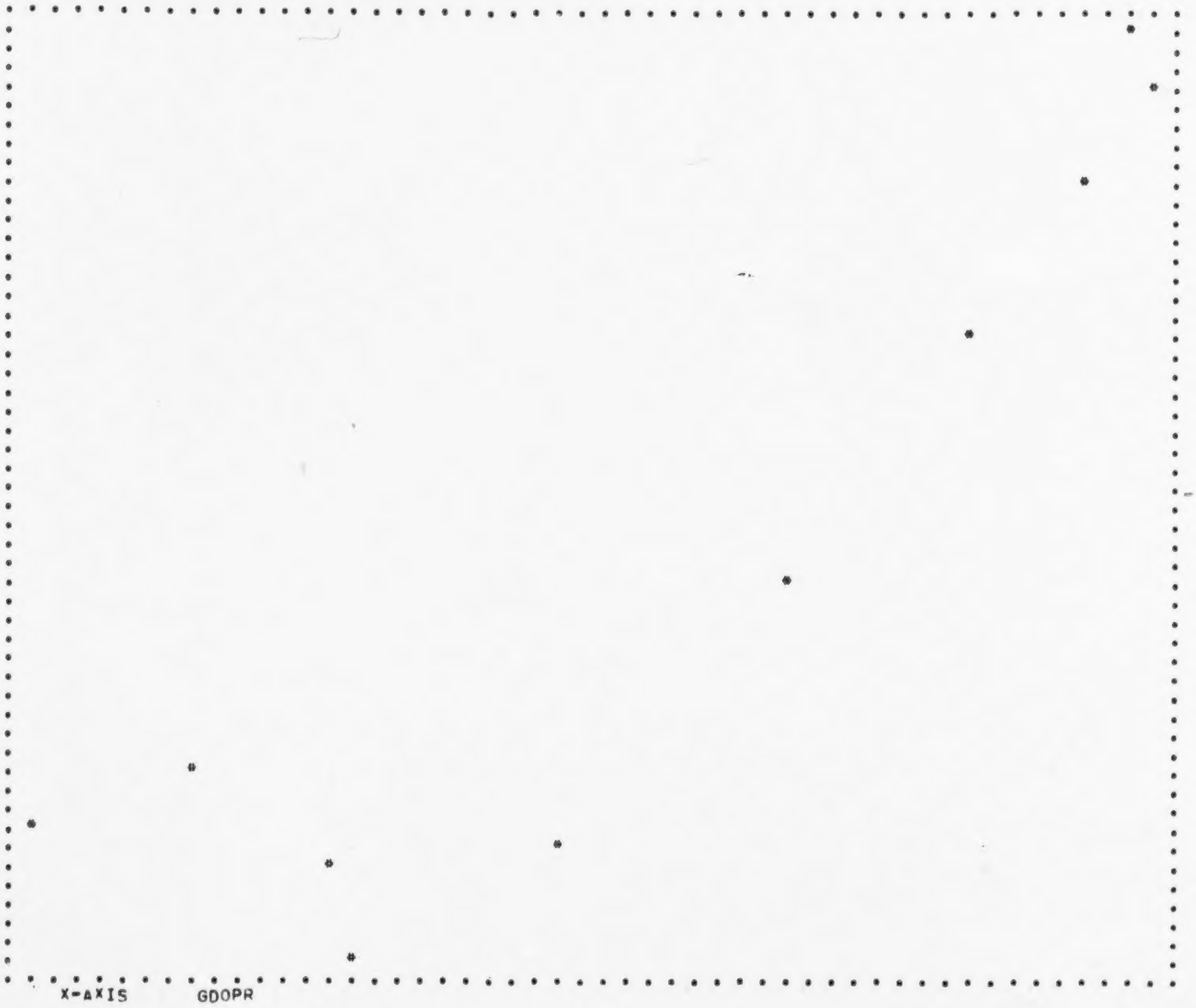
continued

751

```
EQUATION NUMBER          2    NUMBER OF OBSERVATIONS   10
 TOTAL NUMBER OF VARIABLES    5       THIS IS JOHNSON PR.PAGE 269. TEST OF PROGRAM

 TWO STAGE LEAST SQUARES

     VARIANCE-COVARIANCE MATRIX OF COEFFICIENT

  2.59799E-03
 -1.08487E-02  5.87310E-02
 -2.89008E+01  1.04108E+02  3.42742E+05

    DEPENDENT VARIABLE     GFCF

     A( GDOPR  )=       .03386220      T=   .664      BETA COEF. =      .00001956     ELASTICITY(ABOUT THE MEAN)=      .42308653
                (       .05097052)

     A( LAG1   )=       .90631927      T=  3.740      BETA COEF. =      .00248890     ELASTICITY(ABOUT THE MEAN)=      .85866952
                (       .24234480)

     A( CONS   )=  -367.09994742      T=  -.627
                (   585.44173162)

    VARIANCE= 7.78781074E+03

    STAND DEV=8.82485736E+01

    R-SQUARED=      .9018          (ADJUSTED FOR D. OF FREEDOM)

    F-TEST(  2,  7)=    42.3399

    R-SQUARED=      .9236          (NOT ADJUSTED FOR DEGREES OF FREEDOM)



    ACTUALS       PREDICTED       RESIDUALS      0/0 ERROR                          RANGE   980.00  TO  1743.97

   1078.000       1047.808          30.192         2.801    1 .     +  *                                                   .
   1123.000       1112.529          10.471          .932    2 .          +*                                                .
   1052.000       1169.025        -117.025       -11.124    3 .      *        +                                            .
    980.000       1106.471        -126.471       -12.905    4 .*         +                                                 .
   1073.000       1061.229          11.771         1.097    5 .      +*                                                    .
   1281.000       1170.202         110.798         8.649    6 .               +       *                                    .
   1474.000       1377.578          96.422         6.542    7 .                               +      *                     .
   1591.000       1562.995          28.005         1.760    8 .                                             + *            .
   1668.000       1677.195          -9.195         -.551    9 .                                                     *+     .
   1709.000       1743.968         -34.968        -2.046   10 .                                                        *  +.
 DURBIN-WATSON D STATISTIC=    .96463
```

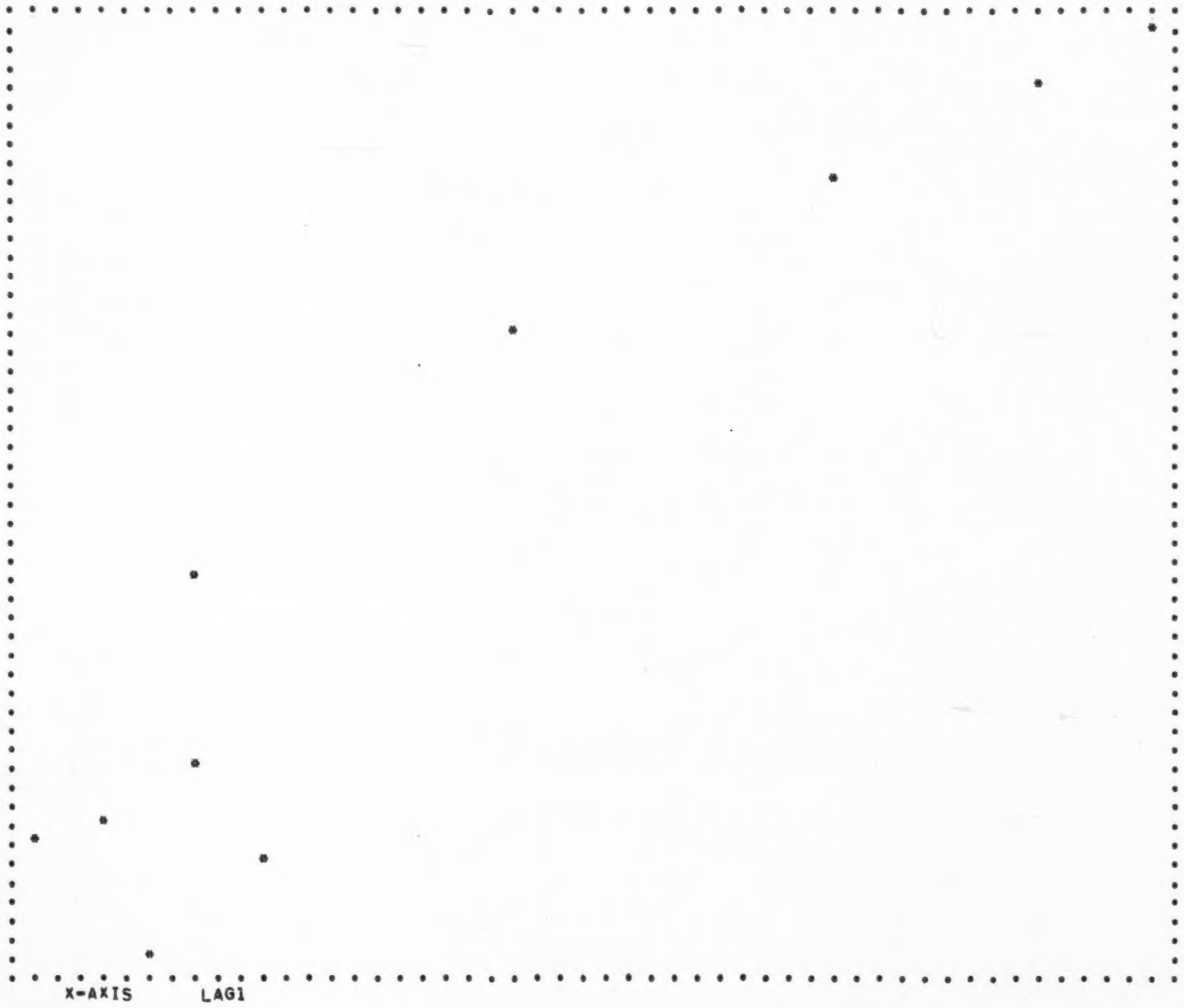
continued

```
SCATTER DIAGRAM WITH VARYING SCALE
Y-AXIS  GFCF    RANGE         980.000        1709.000
X-AXIS  GDOPR   RANGE       14377.000       17788.000
CODE OF SYMBOL   *=1   +=2   $=OVER 4
```



continued

```
SCATTER DIAGRAM WITH VARYING SCALE
Y-AXIS  BFCF    RANGE          980.000        1709.000
X-AXIS  LAG1    RANGE          980.000        1668.000
CODE OF SYMBOL   *=1   +=2   $=OVER 4
```



754

COST ESTIMATE

For the job listed on the Sample Output, the central processor time was 1.144 seconds. At the current rate for the University of Minnesota ($0.20/sec.), the computer time cost $0.23 plus a charge for output supplies.

Charge to user = computer costs + postage + network overhead

= $0.57 + postage + network overhead

CONTENTS—UMST630

pages	
1– 4	Identification & Abstract
5–11	User Instructions
13–22	I/O
23	Cost—Contents

755